

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number  
**WO 01/25406 A1**

(51) International Patent Classification<sup>7</sup>: C12N 9/00,  
9/24, 15/09, 1/20, C07K 16/00

(21) International Application Number: PCT/US00/27210

(22) International Filing Date: 2 October 2000 (02.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/157,341 1 October 1999 (01.10.1999) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant (*for all designated States except US*): UNIVERSITY OF VICTORIA INNOVATION & DEVELOPMENT CORPORATION [CA/CA]; Box 3075 R-Hut, McKenzie Avenue, Victoria, British Columbia V8W 3W2 (CA).

**Published:**

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): HINTZ, William, E. [CA/CA]; 5601 Matterhorn Court, Victoria, British Columbia V9E 2G1 (CA). EADES, Caleb, Joshua [US/CA]; 852 Leslie Drive, Victoria, British Columbia V8X 2Y4 (CA).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(74) Agent: DE GRANDIS, Paula, A.; Klarquist, Sparkman, Campbell, Leigh & Winston, LLP, Suite 1600, One World Trade Center, 121 S.W. Salmon Street, Portland, OR 97204 (US).

(54) Title: MANNOSIDASES AND METHODS FOR USING SAME

(57) Abstract: Mannosidase enzymes and use of such enzymes to alter the glycosylation patterns of macromolecules are disclosed. Also disclosed are the nucleic acid sequences encoding the mannosidase enzymes.

WO 01/25406 A1

## MANNOSIDASES AND METHODS FOR USING SAME

### Field of the Invention

The invention relates to mannosidases and methods of using mannosidases  
5 for altering the glycosylation pattern of macromolecules.

### Background

The N-linked protein glycosylation pathways are fairly well characterized in  
higher eukaryotes (Kornfeld and Kornfeld, *Ann. Rev. Biochem.* 54:631-664, 1985),  
10 however, less is known about such pathways in lower eukaryotes. What is known  
about protein glycosylation in lower eukaryotes has come largely from studies of the  
yeast *Saccharomyces cerevisiae* and may not be sufficient to describe N-linked  
glycosylation in other lower eukaryotes, such as filamentous fungi.

Evolutionary studies suggest that the filamentous ascomycetes diverged from  
15 yeasts from 400 million years ago (Berbee and Taylor, *Mol. Biol. Evol.* 9:278-284,  
1992) to 1 billion years ago, the latter being about the time the fungal branch split  
from plants and animals (Doolittle et al., *Science* 271:470-477, 1996). Filamentous  
fungi produce N-glycan structures that appear to be different than those produced in  
yeast, suggesting a different mode of formation of these structures than in yeast.  
20 Studies have shown the presence in filamentous fungi of N-glycan structures  
containing five mannose units (Man<sub>5</sub>GlcNAc<sub>2</sub>), suggesting processing of a  
Man<sub>5</sub>GlcNAc<sub>2</sub> precursor (Mares et al., *Eur. J. Biochem.* 245:617-625, 1997; Chiba  
et al., *Curr. Microbiol.* 27:281-288, 1993). The mechanisms of synthesis of high  
mannose N-glycans in filamentous fungi seem to differ from corresponding  
25 mechanisms in yeast, and may be more similar to processes in higher eukaryotes.  
Full characterization of the N-glycosylation pathways in these organisms is very  
important, since protein glycosylation plays an integral role in processes such as  
pathogenicity and protein secretion.

The early steps of the asparagine-linked (N-linked) protein glycosylation  
30 pathways are similar in higher and lower eukaryotes (see reviews in Kornfeld and  
Kornfeld, *Ann. Rev. Biochem.* 54:631-664, 1985; Moreman et al., *Glycobiol.* 4:113-  
125, 1994; Herscovics and Orlean, *FASEB J.* 7:540-550, 1993; Herscovics, *Biochim.*

*Biophys. Acta* 1426:275-285, 1999). Initially, an oligosaccharide precursor consisting of three glucose, nine mannose, and two N-acetylglucosamine molecules ( $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$ ) is co-translationally transferred to the newly synthesized polypeptide in the endoplasmic reticulum (ER). In the ER,  $\alpha$ -glucosidases I and II first remove the three glucose molecules. In the ER and/or Golgi apparatus,  $\alpha$ -mannosidase(s) then remove one or more of the mannose residues from the precursor. In higher eukaryotes (i.e., mammals),  $\alpha$ -1,2-mannosidases remove a total of four mannose residues, yielding  $\text{Man}_5\text{GlcNAc}_2$  which is the precursor for complex, hybrid, and high-mannose N-glycans. In the yeast *S. cerevisiae*, however, an ER-specific mannosidase,  $\alpha$ -1,2 mannosidase, removes only a single mannose residue, producing  $\text{Man}_8\text{GlcNAc}_2$  (Herscovics, *Biochim. Biophys. Acta* 1426:275-285, 1999).

Subsequent steps in the pathways in higher and lower eukaryotes are quite different. In higher eukaryotes, following the addition of a single GlcNAc to  $\text{Man}_5\text{GlcNAc}_2$  by GlcNAc transferase I (GnT I), mannosidase II removes two additional mannose groups, producing  $\text{GlcNAcMan}_3\text{GlcNAc}_2$ . Various transferases, such as GnT II, fucosyl transferase, galactosyl transferase, and sialyl transferase, assemble the oligosaccharide into its final structure. In higher eukaryotes a variety of different carbohydrate units can thus be attached to a common precursor to form an array of distinct N-glycans. In *S. cerevisiae*, after the removal of a single mannose, various mannosyltransferases then add additional mannose units to  $\text{Man}_8\text{GlcNAc}_2$  to form large high-mannose N-glycans containing up to 13 mannose units, and even larger mannan outer chains containing up to 200 mannose residues (Herscovics and Orlean, *FASEB J.* 7:540-550, 1993). The removal of mannose residues from the glycan chain during the initial stages of processing appears to be significantly different for higher and lower eukaryotes, resulting in quite different N-glycan structures for these organisms. Based on the evolutionary history of filamentous fungi, their N-glycan processing is likely to contain elements of the pathways in both mammals and yeast. Filamentous fungi do not produce complex N-glycans because such fungi lack the further processive transferases; however, the initial oligosaccharide precursor is trimmed to  $\text{Man}_5\text{GlcNAc}_2$ .

The  $\alpha$ -mannosidases have been classified previously into two independently derived groups, Class 1 and Class 2, based on biochemical properties, substrate specificity, inhibitor profiles, and sequence alignments, (Daniel et al., *Glycobiol.* 4:551-566, 1994; Moreman et al., *Glycobiol.* 4:113-125, 1994; Eades et al.,  
5 *Glycobiol.* 8:17-33, 1998). The first group contains the  $\alpha$ -1,2-mannosidases found in the ER and the Golgi apparatus, including the ER Man<sub>9</sub>-mannosidase, ER endomannosidase, and the Golgi mannosidase I. The second group of  $\alpha$ -mannosidases is more heterogeneous and contains the lysosomal mannosidases, the Golgi mannosidase II, and a distantly related group of enzymes, including the rat  
10 ER/cytosolic mannosidase (Bischoff et al., *J. Biological Chem.* 28:17110-17117, 1990), yeast vacuolar mannosidase (Yoshihisa and Anraku, *Biochem. Biophys. Res. Comm.* 163:908-915, 1989), and the *A. nidulans* Class 2 mannosidase (Eades et al., *Glycobiol.* 8:17-33, 1998).

## 15 Summary of the Invention

The invention stems from the discovery of the amino acid sequences and corresponding nucleic acid sequences for three mannosidases. These mannosidases are particularly useful for altering the glycosylation patterns of macromolecules such as proteins. Glycosylation affects many properties of a glycoprotein, including  
20 protein folding, protease resistance, intercellular trafficking, compartmentalization, secretion, inter- and intra-molecular associations, intermolecular affinities, tissue targeting and biological half-life. Glycosylation patterns may also significantly alter the biological activity, solubility, clearance, intermolecular aggregation, and antigenicity, especially for those proteins having therapeutic utility. Thus, the  
25 present invention enables glycoproteins to be engineered to be more effectively used and produced.

One aspect of the present invention provides the amino acid sequences of three mannosidase proteins. The invention also provides variants of the disclosed mannosidase proteins. These variants can differ from the disclosed sequences by  
30 one or more conservative amino acid substitutions. Additionally, the invention provides variants of the disclosed mannosidase proteins having at least 60% sequence identity to the disclosed mannosidase amino acid sequences.

According to another aspect of the invention, mannosidase-specific binding agents are provided. These specific binding agents bind substantially only the mannosidases and mannosidase variants described above.

According to another aspect of the invention, respective nucleic acid sequences are provided that encode the mannosidases summarized above. These nucleic acid sequences can be operably linked to control sequences and incorporated into any of various vectors. The resulting recombinant vectors are useful for transforming any of various host cells. Once transformed, the host cell can produce the mannosidases of the present invention. Host cells can be obtained from fungi, plants, bacteria, animals, yeast, and insects.

According to yet another aspect of the invention, methods are provided for altering the glycosylation pattern of target proteins using one or more of the mannosidases of the present invention. The mannosidase can be placed in contact with a protein and allowed to alter the glycosylation pattern of the protein. The methods can be practiced either *in vivo* by creating a transgenic host cell that over-expresses or under-expresses one or more of the mannosidases, or *in vitro* by first isolating one or more of the mannosidases of the present invention, and then contacting the isolated mannosidase with a target protein.

According to yet another aspect of the invention, isolated nucleic acid molecules and amino acid molecules are provided that have, by way of example, at least 15, 20, 30, 40, or 50 contiguous nucleotides or amino acid residues with the sequences shown in SEQ ID NOS: 1, 4, and 17; or the complementary strands thereof, or SEQ ID NOS: 3, 6, or 18, respectively.

#### Brief Description of the Drawings

Figures 1A – 1J show a DNA sequence of *A. nidulans* mannosidase 1A. The derived amino acid sequence is indicated by single-letter designations. The hydrophobic transmembrane region is indicated by a dotted underline. Upstream elements are underlined. The intron is indicated in bold type and consensus splice sites are underlined. The stop codon is indicated by a triple asterisk (\*\*\*)

Figures 2A-2G show a DNA sequence of *A. nidulans* mannosidase 1B. The derived amino acid sequence is indicated by single-letter designations. Upstream

elements are underlined. The introns are indicated in bold type and consensus splice sites are underlined. The stop codon is indicated by a triple asterisk (\*\*\*)

**Figure 3** shows a Kyte-Doolittle hydropathy plot of the predicted amino acid sequences of (A) *A. nidulans* mannosidase 1A and (B) *A. nidulans* mannosidase 1B.

5 The vertical axis is the hydrophobicity of a given region of the protein (Kyte and Doolittle, *J. Mol. Biol.* **157**:105-132, 1982) with positive values representing hydrophobic regions of the protein, and negative values representing hydrophilic regions of the protein.

**Figure 4** shows a sequence-similarity matrix generated from the ClustalW (Thompson et al., *Nucl. Acids Research* **22**:4673-4680, 1994) sequence alignment of 13 Class 1  $\alpha$ -mannosidases and the pairwise percent similarity of the amino acid sequences for all possible pairs of sequences.

**Figure 5** shows a dendrogram showing sequence relationships of mannosidases. The dendrogram was generated from the ClustalW sequence alignment. Three major groups of related enzymes are shown in shaded boxes.

**Figures 6A and 6B** show the amino acid and nucleic acid sequences of mannosidase 1C.

### Sequence Listings

The nucleic and amino acid sequences listed in the accompanying sequence listing are shown using standard letter abbreviations for nucleotide bases, and three-letter code for amino acids. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand.

- 25 **SEQ ID NO: 1** is the nucleic acid sequence of the mannosidase 1A gene.
- SEQ ID NO: 2** is the nucleic acid sequence of the mannosidase 1A cDNA.
- SEQ ID NO: 3** is the deduced amino acid sequence of mannosidase 1A.
- SEQ ID NO: 4** is the nucleic acid sequence of the mannosidase 1B gene.
- SEQ ID NO: 5** is the nucleic acid sequence of the mannosidase 1B cDNA.
- SEQ ID NO: 6** is the deduced amino acid sequence of mannosidase 1B.
- 30 **SEQ ID NO: 7** is a 5'-splice site sequence of the mannosidase 1A gene.
- SEQ ID NO: 8** is a 5'-splice site consensus sequence of filamentous fungi.
- SEQ ID NO: 9** is an internal lariat sequence of the mannosidase 1A gene.

SEQ ID NO: 10 is an internal lariat consensus sequence of filamentous fungi.

SEQ ID NO: 11 is a 5'-splice site sequence of the mannosidase 1B gene.

SEQ ID NO: 12 is an internal lariat sequence of the mannosidase 1B gene.

5 SEQ ID NO: 13 is the amino acid sequence used to generate a forward primer.

SEQ ID NO: 14 is the amino acid sequence used to generate a reverse primer.

SEQ ID NO: 15 is a PCR primer useful for identifying mannosidases.

10 SEQ ID NO: 16 is a PCR primer useful for identifying mannosidases.

SEQ ID NO: 17 is the nucleic acid sequence of the mannosidase 1C gene.

SEQ ID NO: 18 is the deduced amino acid sequence of mannosidase 1C.

SEQ ID NO: 19 is the deduced amino acid sequence of a consensus splice site.

15

### Detailed Description

#### I. Definitions

“Sequence Identity.” The similarity between two nucleic acid sequences or between two amino acid sequences is expressed in terms of the level of sequence identity shared between the sequences. Sequence identity is typically expressed in terms of percentage identity; the higher the percentage, the more similar the two sequences are.

Methods for aligning sequences for comparison purposes are well known in the art. Various programs and alignment algorithms are described in: Smith & Waterman, *Adv. Appl. Math.* 2:482, 1981; Needleman & Wunsch, *J. Mol. Biol.* 48:443, 1970; Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; Higgins & Sharp, *Gene* 73:237-244, 1988; Higgins & Sharp, *CABIOS* 5:151-153, 1989; Corpet et al., *Nucleic Acids Research* 16:10881-10890, 1988; Huang, et al., *Computer Applications in the Biosciences* 8:155-165, 1992; and Pearson et al., *Methods in Molecular Biology* 24:307-331, 1994. Altschul et al., *J. Mol. Biol.*, 215:403-410, 1990, presents a detailed consideration of sequence alignment methods and homology calculations.



The NCBI Basic Local Alignment Search Tool (BLAST™, Altschul et al., *J. Mol. Biol.*, 215:403-410, 1990) is available from several sources, including the National Center for Biotechnology Information (NCBI, Bethesda, MD) and on the Internet, for use in connection with the sequence-analysis programs blastp, blastn, blastx, tblastn and tblastx. BLAST™ can be accessed on the NCBI website. A description of how to determine sequence identity using this program is available on the internet at the NCBI website.

For comparisons of amino acid sequences of greater than about 30 amino acids, the "Blast 2 sequences" function in the BLAST™ program is employed using the default BLOSUM62 matrix set to default parameters, (gap existence cost of 11, and a per-residue gap cost of 1). When aligning short peptides (fewer than about 30 amino acids), the alignment should be performed using the Blast 2 sequences function, employing the PAM30 matrix set to default parameters (open gap 9, extension gap 1 penalties). Proteins with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 45%, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% sequence identity.

**"Substantial similarity"** A first nucleic acid is "substantially similar" to a second nucleic acid if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 60%, 75%, 80%, 85%, 90% or 95% of the nucleotide bases. Sequence similarity can be determined by comparing the nucleotide sequences of two nucleic acids using the BLAST™ sequence analysis software (blastn) available from The National Center for Biotechnology Information. Such comparisons may be made using the software set to default settings (expect = 10, filter = default, descriptions = 500 pairwise, alignments = 500, alignment view = standard, gap existence cost = 11, per residue existence = 1, per residue gap cost = 0.85). Similarly, a first polypeptide is substantially similar to a second polypeptide if it shows sequence identity of at least about 75%-90% or greater when optically aligned and compared using BLAST™ software (blastp) using default settings.

**“Specific Binding Agent.”** A “specific binding agent” is an agent that is capable of specifically binding to the mannosidases of the present invention, and may include polyclonal antibodies, monoclonal antibodies (including humanized  
5 monoclonal antibodies) and fragments of monoclonal antibodies such as Fab, F(ab')<sub>2</sub> and Fv fragments, as well as any other agent capable of specifically binding to the epitopes on the proteins.

**“Operably linked.”** A first nucleic acid sequence is “operably linked” with  
10 a second nucleic acid sequence whenever the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA sequences are contiguous and, where necessary to join two protein-coding regions,  
15 in the same reading frame.

**“Isolated.”** An “isolated” biological component (such as a nucleic acid or protein or organelle) is a component that has been substantially separated or purified away from other biological components in the cell of the organism in which the  
20 component naturally occurs, i.e., other chromosomal and extra-chromosomal DNA, RNA, proteins, and organelles. Nucleic acids and proteins that have been “isolated” include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell, as well as chemically synthesized nucleic acids.

25  
**“Recombinant.”** A “recombinant” nucleic acid is one having a sequence that is not naturally occurring or has a sequence made by an artificial combination of two otherwise-separated, shorter sequences. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial  
30 manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques.

**“Transformed.”** A “transformed” cell is a cell into which a nucleic acid molecule has been introduced by molecular biology techniques. As used herein, the term “transformation” encompasses all techniques by which a nucleic acid molecule might be introduced into such a cell, including transfection with a viral vector,  
5 transformation with a plasmid vector, and introduction of naked DNA by electroporation, lipofection, and particle gun acceleration.

**“Vector.”** A “vector” is a nucleic acid molecule as introduced into a host cell, thereby producing a transformed host cell. A vector may include nucleic acid  
10 sequences, such as an origin of replication, that permit the vector to replicate in a host cell. A vector may also include one or more selectable marker genes and other genetic elements known in the art.

**“Glycosylation pattern.”** The “glycosylation pattern” is the characteristic  
15 structure (including branch structure), number, or location of oligosaccharide structures associated with a macromolecule, such as a protein.

**“DNA construct.”** The term “DNA construct” is intended to denote any nucleic acid molecule of cDNA, genomic DNA, synthetic DNA or RNA origin. The  
20 term “construct” is intended to denote a nucleic acid segment that may be single- or double-stranded, and that may be based on a complete or partial naturally occurring nucleotide sequence encoding one or more of the mannosidase genes of the present invention. It is understood that such nucleotide sequences include intentionally manipulated nucleotide sequences, e.g., subjected to site-directed mutagenesis, and  
25 sequences that are degenerate as a result of the genetic code. All degenerate nucleotide sequences are included within the scope of the invention so long as the mannosidase enzyme encoded by the nucleotide sequence maintains the ability to hydrolytically remove terminal mannoside residues.

**“Mannosidase activity.”** The phrase “mannosidase activity” describes the  
30 enzymatic catalysis of the hydrolytic removal of terminal mannoside residues.

“Probes and primers.” Nucleic acid probes and primers may be readily prepared based on the nucleic acid sequences provided by this invention. A “probe” comprises an isolated nucleic acid sequence attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed, e.g., in Sambrook et al. (eds.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; and Ausubel et al. (ed.) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York (with periodic updates), 1987.

“Primers” are short nucleic acids, preferably DNA oligonucleotides 15 nucleotides or more in length, that are annealed to a complementary target DNA strand by nucleic acid hybridization to form a hybrid between the primer and the target DNA strand, then extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR) or other nucleic-acid amplification methods known in the art.

As noted, probes and primers are preferably 15 nucleotides or more in length, but, to enhance specificity, probes and primers of 20 or more nucleotides may be preferred.

Methods for preparing and using probes and primers are described, for example, in Sambrook et al. (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; Ausubel et al. (ed.), *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York (with periodic updates), 1987; and Innis et al., *PCR Protocols: A Guide to Methods and Applications*, Academic Press: San Diego, 1990. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge, MA). One of skill in the art will appreciate that the specificity of a particular probe or primer increases with the length of the probe or primer. For example, a primer comprising 20 consecutive nucleotides will anneal to a target with a higher specificity than a

corresponding primer of only 15 nucleotides. Thus, in order to obtain greater specificity, probes and primers may be selected that comprise, by way of example, 10, 20, 25, 30, 35, 40, 50 or more consecutive nucleotides.

## 5 II. Methods of Producing Mannosidase 1A, 1B, and 1C

### A. Cloning Nucleic Acid Sequences Encoding Mannosidase

Provided with the nucleic acid sequences of the genes encoding the mannosidases 1A, 1B, and 1C (SEQ ID NOS: 1, 4, and 17, respectively), one of  
10 ordinary skill in the art will appreciate that several different methods can be used to isolate the genes and the cDNAs encoding the corresponding mannosidases. One example of such a method is the polymerase chain reaction (PCR) (U.S. Pat. No. 4,683,202 to Mullis; and Saiki et al., *Science* 239:487-491, 1988).

When using PCR to isolate a sequence encoding the gene, a primer can be  
15 designed that targets the extreme 5' end of the sequence, and a second primer can be designed that targets the extreme 3' end of the sequence. For example the 5' primer (5'-GGYGGYCTNGGYGARTCNTTCTACGAGTA-3'; SEQ ID NO: 15) and the 3' primer (5'-GTANAGGTACTTNAGNGTCTCNGCNAGRHAGAA-3'; SEQ ID NO: 16) are used in a PCR (polymerase chain reaction) procedure to generate  
20 multiple copies of the gene. The copies are isolated by separation on an agarose gel. The fragment of interest is then removed from the gel, and ligated into an appropriate vector.

Alternatively, the gene can be created by engineering synthetic strands of DNA that partially overlap each other (Beaucage & Caruthers, *Tetrahedron Letters*  
25 22:1859-1869, 1981; Matthes et al., *Embo. J.* 3:801-805, 1984). The synthetic strands are annealed and a DNA polymerase is used to fill in the single-stranded regions. The resulting synthetic double-stranded DNA molecule can be cloned into a vector.

For use as primers and probes, nucleic acid sequences can contain at least 15  
30 contiguous nucleic acid molecules of either of the sequences shown in Seq. ID. No. 1 and SEQ ID NO: 4, or either of the complementary strands of the molecules shown in SEQ ID NO: 1 and SEQ ID NO: 4. The nucleic acid sequences are useful

for performing hybridization protocols, such as Northern blots or Southern blots as described in Sambrook et al., (eds.), *Molecular Cloning, A Laboratory Manual*, 2d ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

These hybridization protocols can be used to identify nucleic acid sequences that are substantially similar to those shown in SEQ ID NOS: 1, 4, or 17. A successful hybridization to such sequences indicates that the analogous nucleic acid sequence hybridizes to the oligonucleotide probe that comprises at least a fragment of the sequences shown in SEQ ID NOS: 1, 4, or 17. Generally hybridization conditions are classified into categories, for example very high stringency, high stringency, and low stringency. The conditions corresponding to these categories are provided below.

Very High Stringency (detects sequences that share 90% sequence identity)

	Hybridization	in	5x	SSC	at	65°C	16 hours
15	Wash twice	in	2x	SSC	at	room temp.	15 minutes each
	Wash twice	in	0.2x	SSC	at	65°C	20 minutes each

High Stringency (detects sequences that share 80% sequence identity or greater)

20	Hybridization	in	3x	SSC	at	65°C	16 hours
	Wash twice	in	2x	SSC	at	room temp.	15 minutes each
	Wash twice	in	0.5x	SSC	at	55°C	20 minutes each

Low Stringency (detects sequences that share greater than 50% sequence identity)

25	Hybridization	in	3x	SSC	at	65°C	16 hours
	Wash twice	in	2x	SSC	at	room temp.	20 minutes each

Mannosidase-encoding nucleic acid sequences according to the invention also encompass mannosidase enzymes that differ in amino acid sequence from the mannosidase 1A, 1B, and 1C sequences of SEQ ID NOS: 1, 4, and 17, and that maintain mannosidase activity. Such proteins may be produced by changing the

cDNA nucleotide sequence of mannosidase 1A, 1B, or 1C, by changing the sequence of the respective genes using standard procedures such as site-directed mutagenesis, or by performing the polymerase chain reaction. The simplest modifications involve substituting one or more amino acids with other amino acids having similar biochemical properties. These so-called "conservative substitutions" usually have minimal impact on the activity of the resultant protein. Table 1 shows amino acids that may be substituted for an original amino acid in a protein and that are regarded as conservative substitutions.

**Table 1**

<b>Original Residue</b>	<b>Conservative Substitutions</b>
ala	ser
arg	lys
asn	gln; his
asp	glu
cys	ser
gln	asn
glu	asp
gly	pro
his	asn; gln
ile	leu; val
leu	ile; val
lys	arg; gln; glu
met	leu; ile
phe	met; leu; tyr
ser	thr
thr	ser
trp	tyr
tyr	trp; phe
val	ile; leu

More substantial changes in enzymatic function or other features may be obtained by selecting substitutions that are less conservative than those in Table 1, i.e., selecting residues that differ more significantly in their effect on maintaining: (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Substitutions that

generally produce the greatest changes in protein properties are those in which: (a) a hydrophilic residue, e.g., seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g., leucyl, isoleucyl, phenylalanyl, valyl, or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an  
5 electropositive side chain, e.g., lysyl, arginyl, or histadyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or (d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine. The effects of these amino acid substitutions or deletions or additions may be assessed for mannosidase protein derivatives by analyzing the ability of the  
10 respective modified polypeptide to catalyze the removal of terminal mannoside residues.

Variant mannosidase cDNA or genes may be produced by standard DNA mutagenesis techniques, for example, M13 primer mutagenesis. Details of these techniques are provided in Sambrook et al. (ed.), *Molecular Cloning: A Laboratory  
15 Manual*, 2nd ed., vol. 1-3, Ch. 15, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, and Ausubel et al. (ed.) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York (with periodic updates), 1987. By the use of such techniques, variants may be created that differ slightly from the mannosidase cDNA or gene sequences specifically disclosed, yet  
20 that still encode a protein having mannosidase activity. DNA molecules and nucleotide sequences that are derivatives of those specifically disclosed herein and that differ from those disclosed by the deletion, addition, or substitution of nucleotides while still encoding a protein having mannosidase activity are comprehended by this invention. In their simplest form, such variants may differ  
25 from the disclosed sequences by alteration of the coding region to fit the codon usage bias of the particular organism into which the molecule is to be introduced.

Alternatively, the coding region may be altered by taking advantage of the degeneracy of the genetic code to alter the coding sequence in such a way that, while the nucleotide sequence is substantially altered, it nevertheless encodes a protein  
30 having an amino acid sequence identical or substantially similar to the mannosidase 1A and 1B sequences specifically disclosed in SEQ ID NOS: 3 and 6. For example, the fourth amino acid residue of the mannosidase 1A gene is alanine and is encoded



in the mannosidase 1A open reading frame (ORF) by the nucleotide codon triplet GCA. Because of the degeneracy of the genetic code, three other nucleotide codon triplets-GCT, GCC and GCG-also code for alanine. Thus, the nucleotide sequence of the mannosidase 1A ORF could be changed at this position to any of these three  
5 codons without affecting the amino acid composition of the encoded protein or the characteristics of the protein. Based upon the degeneracy of the genetic code, variant DNA molecules may be derived from the cDNA and gene sequences disclosed herein using standard DNA mutagenesis techniques as described above, or by synthesis of modified DNA sequences. Thus, this invention also encompasses  
10 nucleic acid sequences that encode either a mannosidase 1A or 1B protein but that vary from the disclosed nucleic acid sequences due to the degeneracy of the genetic code.

#### B. Vectors

15 The choice of expression vector depends in part on the type of host cell that will be used for the expression of the mannosidase enzyme. The promoter sequence used must be recognizable by the enzymes responsible for translation within the host cell. One of skill in the art will appreciate that there are a number of regulatory sequences known in the art that function in conjunction with bacterial host cells,  
20 plant host cells, insect host cells, yeast host cells, and fungal host cells.

Examples of constitutive plant promoters that may be useful for expressing the cDNA include: the cauliflower mosaic virus (CaMV) 35S promoter, which confers constitutive, high-level expression in most plant tissues (see, e.g., Odel et al., *Nature* 313:810, 1985), nopaline synthase promoter (An et al., *Plant Physiol.*  
25 88:547, 1988); and the octopine synthase promoter (Fromm et al., *Plant Cell* 1:977, 1989).

Examples of suitable promoters for use in filamentous fungus host cells are, for instance, the ADH3 promoter (McKnight et al., *Embo. J.* 4:2093-2099, 1985) or the *tpiA* promoter (Blattner et al., *Science* 27:1453-1474, 1997). Examples of other  
30 useful promoters are those derived from the genes encoding *A. oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *A. niger* neutral  $\alpha$ -amylase, *A. niger* acid-stable  $\alpha$ -amylase, *A. niger* or *A. awamori* glucoamylase (gluA),

*Rhizomucor miehei* lipase, *A. oryzae* alkaline protease, *A. oryzae* triose phosphate isomerase, or *A. nidulans* acetamidase. It is also possible that the native mannosidase promoter could be used to express the enzyme. Similarly, one of ordinary skill in the art will appreciate that there are several promoters available that can be used in insect expression systems (e.g., baculovirus) and bacterial expression systems.

To direct the enzyme into the secretory pathway of the host cell, a secretory signal sequence (also known as a leader sequence, "prepro" sequence or "pre" sequence) may be provided in the vector. The secretory signal sequence is joined to the DNA sequence encoding the enzyme in the correct reading frame. Secretory signal sequences are commonly positioned 5' to the DNA sequence encoding the enzyme. The secretory signal sequence may be the sequences normally associated with the enzyme or from a gene encoding another secreted protein.

The vector will also most likely contain a selectable marker. The selectable marker allows host cells that have been successfully transformed with the construct of interest to be identified. Selectable markers for use in plant cells are, for example Bost<sup>r</sup>, Kan<sup>r</sup>, and various other herbicide resistance genes. Selectable markers for use in mammalian cells are, for example, Amp<sup>r</sup> and Kan<sup>r</sup>. Selectable markers for use in filamentous fungi include, for example, amdS, pyrG, argB, niaD and sC.

20

### C. Transformation

The DNA construct of the invention may be either homologous or heterologous to the host in question. If homologous to the host cell, i.e., produced by the host cell in nature, the construct typically will be operably connected to another promoter sequence or, if applicable, another secretory signal sequence and/or terminator sequence than in its natural environment. In this context, the term "homologous" is intended to include a cDNA sequence encoding the mannosidase 1A, 1B, or 1C native to the host *A. nidulans*. The term "heterologous" is intended to include a DNA sequence not expressed by the host cell in nature. Thus, the DNA sequence may be from another organism, or it may be a synthetic sequence.

30

The host cell of the invention, into which the DNA construct or the recombinant expression vector of the invention is to be introduced, may be any cell

capable of producing the disclosed mannosidase enzymes. Such cells include bacteria cells, yeast cells, fungal cells, insect cells, and higher eukaryotic cells.

Various methods of introducing the DNA construct into host cells are well known in the art. For example, in some species, the Ti plasmid of *A. tumefaciens* can be used to transform host cells (Gouka et al., *Nature Biotechnology* 6:598-602, 1999), the host cell can also be transformed using gene blasting techniques and standard chemical treatments.

#### D. Production

The production of recombinant mannosidase 1A, 1B, or 1C may be accomplished using any of various transformed hosts, such as bacteria, plants, mammalian cell culture, whole mammals, insects, yeast, and fungi. As mentioned above, the selection of suitable promoter sequences and other regulatory elements will be specific to the host organism that will be used to produce the protein. Additionally, within each group of potential hosts there are several species that can potentially be used to produce the protein. For example, filamentous fungi, e.g., *Aspergillus* spp., *Neurospora* spp., *Fusarium* spp. or *Trichoderma* spp., can be used to produce mannosidase. More specifically, individual strains of *Aspergillus*, such as *A. oryzae*, *A. nidulans*, or *A. niger*, can be used. The use of *Aspergillus* spp. for the expression of proteins is described in, e.g., EP 272,277 and EP 230,023. Transformation techniques specific for the transformation of *F. oxysporum* have been described by Malardier et al., *Gene* 78:147-156, 1988.

Various yeast strains and yeast-derived vectors are commonly used for the expression of heterologous proteins. For instance, *Pichia pastoris* expression systems, obtained from Invitrogen (San Diego, California), may be used to practice the present invention. Such systems include suitable *Pichia pastoris* strains, vectors, reagents, transformants, sequencing primers, and media. Available strains include KM71H a prototrophic strain, SMD1168H a prototrophic strain, and SMD1168 a pep4 mutant strain (Invitrogen Product Catalogue, 1998, Invitrogen, Carlsbad CA).

Non-yeast eukaryotic vectors may be used with equal facility for expression of proteins encoded by modified nucleotides according to the invention. Mammalian vector/host cell systems containing genetic and cellular control

elements capable of carrying out transcription, translation, and post-translational modification are well known in the art. Examples of such systems are the well-known Baculovirus system, the ecdysone-inducible mammalian expression system that uses regulatory elements from *Drosophila melanogaster* to allow control of gene expression, and the Sindbis viral-expression system that allows high-level expression in a variety of mammalian cell lines, all of which are available from Invitrogen.

The cloned expression vector may be transformed into any of various cell types for expression of the cloned nucleotide. Many different types of cells may be used to express modified nucleic acid molecules. Examples include cells of yeasts, fungi, insects, mammals, and plants, including transformed and non-transformed cells. For instance, common mammalian cells that could be used for the invention include HeLa cells, SW-527 cells (ATCC deposit #7940), WISH cells (ATCC deposit #CCL-25), Daudi cells (ATCC deposit #CCL-213), Mandin-Darby bovine kidney cells (ATCC deposit #CCL-22), and Chinese hamster ovary (CHO) cells (ATCC deposit #CRL-2092). Common yeast cells include *Pichia pastoris* (ATCC deposit #201178) and *Saccharomyces cerevisiae* (ATCC deposit #46024). Insect cells include cells from *Drosophila melanogaster* (ATCC deposit #CRL-10191), the cotton bollworm (ATCC deposit #CRL-9281) and from *Trichoplusia ni* egg cell homoflagellates. Fish cells that may be used include those from rainbow trout (ATCC deposit #CLL-55), salmon (ATCC deposit #CRL-1681), and zebrafish (ATCC deposit #CRL-2147). Amphibian cells that may be used include those of the bullfrog, *Rana catesbeiana* (ATCC deposit #CLL-41). Reptile cells that may be used include those from Russell's viper (ATCC deposit #CCL-140). Plant cells that could be used include *Chlamydomonas* cells (ATCC deposit #30485), *Arabidopsis* cells (ATCC deposit #54069) and tomato plant cells (ATCC deposit #54003). Many of these cell types are commonly used and are available from the ATCC as well as from commercial suppliers such as Pharmacia (Uppsala, Sweden), and Invitrogen (San Diego, California).

Expressed protein may be accumulated within a cell or may be secreted from the cell. Such expressed protein may then be collected and purified. This protein

may then be characterized for activity and heat stability and may be used to practice any of the various methods according to the invention.

#### **E. Isolation and Purification of Mannosidase 1A, 1B, and 1C**

5 The mannosidases of the present invention can be isolated and purified from either transgenic host cells, or from wild-type host cells. The purification of the mannosidase enzymes of the present invention is achieved by first isolating the enzymes from the other cellular components and then purifying the enzymes.

##### **1. Separation of Enzymes from Cellular Components**

10 A variety of methods are known in the art for separating enzymes from other cellular components (complexes). Such methods can involve treatment with harsh chemicals and/or severe temperatures. Typically, protein complexes can be disrupted through the use of reducing agents, denaturants, freeze/thaw cycles, mechanical shearing, decompression/compression, sonication, agitation, and/or  
15 increased temperatures. Protein complexes also can be disrupted using a combination of such techniques.

Reducing agents are capable of donating hydrogen atoms, and thus serve to cleave disulfide bonds. Reducing agents particularly disrupt disulfide bonds that  
20 link two proteins or portions of the same protein together. Commonly used reducing reagents are  $\beta$ -mercaptoethanol, dithiothreitol (DTT), and trialkyl phosphines. The ability of a reducing agent to disrupt a complex is increased by increasing the temperature at which a mixture of the reducing agent and complex is incubated.

Denaturants serve to relax the conformational structure of proteins. Any of  
25 various denaturants can be employed to separate a protein from other cellular components. Examples of denaturants useful in conjunction with ion-exchange columns are urea and formamide. However, denaturants such as guanidine hydrochloride or guanidine thiocyanate may be useful for separating proteins from other cellular components in methods not involving an ion-exchange column. The  
30 ability of a particular denaturant to relax a protein is enhanced by increasing the temperature at which the sample is incubated.

## 2. Purification of Enzymes from Cellular Components

Purification of enzymes from a crude lysate can be achieved by successive rounds of chromatography. In between each round of chromatography, the sample is assayed for enzymatic activity. Examples 7 and 8, below, describe activity assays that can be used to monitor the recovery of mannosidase activity, respectively. Typically, purification involves a multi-step procedure that includes well-known chromatographic techniques (Robyt and White, *Biochemical Techniques Theory and Practice*, Waveland Press, Inc., 1990). In such multi-step procedures, the enzymes are separated by exploiting their different physical characteristics. The following discussion provides a broad description of various chromatography techniques that can be used in either a single-step process or in a multi-step process. A single type of chromatography can be repeatedly used (rather than changing the chromatography step each time) to purify the enzymes of the present invention.

### i. Column Chromatography

One method of isolating an enzyme according to the present invention is adsorption chromatography. This method exploits a protein's differential affinity for the medium in a column, compared to the protein's affinity for the eluting solvent. An example of a suitable medium for use in the column is hydroxyapatite (crystalline calcium phosphate). Hydroxyapatite tends to adsorb acidic proteins that can be subsequently eluted with phosphate ions (phosphate ions have a high affinity for the calcium ions present in the hydroxyapatite).

Ion-exchange chromatography (a variation of adsorption chromatography) can also be used to isolate enzymes according to the present invention. In ion-exchange chromatography, a solid adsorbent is used that has charged groups chemically linked to an inert solid. An ionic charge on an enzyme molecule causes the molecule to attach to an oppositely charged group on the solid support. The enzyme is subsequently released from the support by passing a solution containing an ion gradient over the solid adsorbent. Examples of solid supports are DEAE-cellulose, DEAE-Sephadex<sup>TM</sup>, DEAE-Bio-Gel<sup>TM</sup>, DEAE-Sepharose<sup>TM</sup>, DEAE-Sepharose<sup>TM</sup>, DEAE-Trisacryl<sup>TM</sup>, Q-Sepharose<sup>TM</sup>, ecteola cellulose, QAE-cellulose,

express ion exchanger Q, PEI-cellulose, and other polystyrene-based anion exchangers (most of these support materials are available from Pharmacia).

Another suitable type of adsorption chromatography is affinity chromatography. This method involves covalently linking to an inert solid support a  
5 ligand having a binding affinity for the subject enzyme. Commonly, the ligand is a specific binding agent that selectively binds the enzyme as the enzyme molecule contacts the solid support. Alternatively, the enzymes can be purified based upon their hydrophobicity. For example, alkyl chains can be linked to the inert support to supply sites for hydrophobic bonding interactions between the support and the  
10 enzyme. The eluting solvent contains a hydrophobic gradient.

High-performance liquid chromatography (HPLC) is yet another suitable method. All of the major classes of chromatographic separations are compatible with this method, for example: adsorption, liquid-liquid partition, ion exchange, exclusion, and affinity chromatography can be used in conjunction with HPLC.  
15 Additionally, HPLC allows for reverse-phase and ion-pair partition. Reverse-phase partition is a relatively quick elution technique using a non-polar stationary phase and a polar mobile phase. Ion-pair partition involves pairing a charged polar substance with its counter-ion to create a less polar species that then flows through the column.

20

## ii. Electrophoresis

Electrophoresis is a well-established technique for the separation and analysis of mixtures by differential migration and separation of molecules in an electric field, based on differences in mobility of the molecule through a support.  
25 Many different forms of electrophoresis have been developed to permit the separation of different classes of compounds. These forms include paper and cellulose acetate electrophoresis, thin-layer electrophoresis, gel electrophoresis, immunoelectrophoresis, and isoelectric focusing. Paper electrophoresis operates best for the separation of protein molecules having a relatively low molecular  
30 weight, and gel electrophoresis is better for use in isolating enzymes with higher molecular weights. A variety of different matrices are available for forming gels, but matrices with a relatively small pore size are most suitable for the separation of

two proteins that have similar physical characteristics. Furthermore, proteins may be separated based upon their molecular size by conducting electrophoresis under dissociating conditions, for example using sodium dodecyl sulfate (SDS). SDS relaxes the protein conformation and masks the ionic charge of the protein, hereby  
5 leaving the relative length of the protein as the principal distinguishing characteristic for purposes of separation from other proteins.

Separation also can be achieved using immunoelectrophoresis, in which proteins are separated on a gel, based upon the relative charge-to-mass ratio and antigenicity of the proteins. The proteins are separated on a gel. An antibody,  
10 specific for the protein of interest, is added to a well created in the gel, and the antibody is allowed to diffuse through the gel. A precipitate forms in regions in which the antibody reacts with the protein.

Finally, the enzymes can be purified by isoelectric focusing (IEF). IEF is a type of electrophoresis in which the protein is placed on a substrate having a pH  
15 gradient. The protein moves under the influence of an applied electrical field until the protein reaches a zone in the pH gradient corresponding to the isoelectric point of the protein.

Specific examples of previously published mannosidase-purification techniques are available in Castellion and Bretthauer, *Biochem. J.*, **324**:951-956,  
20 1997; and Eades et al., *Glycobiology* **8**:17-33, 1998.

### **III. Mannosidase Specific Antibodies**

Antibodies to the mannosidase enzymes of the present invention may be useful for purification of the enzymes. The provision of the amino acid sequences of  
25 the mannosidase 1A, 1B, and 1C enzymes allows for the production of specific antibody-based binding agents to these enzymes.

#### **A. Production of an Antibodies to Mannosidase 1A, 1B and/or 1C**

Monoclonal or polyclonal antibodies may be produced to either of the  
30 mannosidase enzymes, portions of the enzymes, or variants thereof. Optimally, antibodies raised against epitopes on these antigens will specifically detect the enzyme. That is, antibodies raised against the enzymes recognize and bind the



enzymes, but do not substantially recognize or bind to other proteins. The determination that an antibody specifically binds to an antigen is made by any one of a number of standard immunoassay methods, for instance, Western blotting, Sambrook et al. (ed.), *Molecular Cloning: A Laboratory Manual*, 2nd ed., vol. 1-3, 5 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

To determine that a given antibody preparation (such as a preparation produced in a mouse against mannosidase 1A) specifically detects a mannosidase by Western blotting, total cellular protein is extracted from fungal cells and electrophoresed on an SDS-polyacrylamide gel. The proteins are then transferred to 10 a membrane (for example, nitrocellulose) by Western blotting, and the antibody preparation is incubated with the membrane. After washing the membrane to remove non-specifically bound antibodies, the presence of specifically bound antibodies is detected by the use of an anti-mouse antibody conjugated to an enzyme such as alkaline phosphatase; application of 5-bromo-4-chloro-3-indolyl 15 phosphate/nitro blue tetrazolium results in the production of a dense blue compound by immuno-localized alkaline phosphatase.

Antibodies that specifically detect mannosidase can be shown, by this technique, to bind substantially only the mannosidase band (having a position on the gel determined by the molecular weight of the mannosidase). Non-specific binding 20 of the antibody to other proteins may occur and may be detectable as a weaker signal on the Western blot (which can be quantified by automated radiography). The non-specific nature of this binding can be recognized by one skilled in the art by the weak signal obtained on the Western blot relative to the strong primary signal arising from the specific anti-mannosidase binding.

25 Antibodies that specifically bind to mannosidase belong to a class of molecules that are referred to herein as "specific binding agents." Specific binding agents that are capable of specifically binding to the mannosidases of the present invention may include polyclonal antibodies, monoclonal antibodies, and fragments of monoclonal antibodies such as Fab, F(ab')<sub>2</sub> and Fv fragments, as well as any 30 other agent capable of specifically binding to one or more epitopes on the proteins.

Substantially pure mannosidase 1A or 1B suitable for use as an immunogen can be isolated from transfected cells, transformed cells, or from wild-type cells.

Concentration of protein in the final preparation is adjusted, for example, by concentration on an Amicon filter device (Charlotte, North Carolina), to the level of a few micrograms per milliliter. Alternatively, peptide fragments of either of the mannosidases may be utilized as immunogens. Such fragments may be chemically synthesized using standard methods, or may be obtained by cleavage of the whole mannosidase enzyme followed by purification of the desired peptide fragments. Peptides as short as three or four amino acids in length are immunogenic when presented to an immune system in the context of a Major Histocompatibility Complex (MHC) molecule, such as MHC class I or MHC class II. Accordingly, peptides comprising at least 3 and preferably at least 4, 5, 6 or more consecutive amino acids of the disclosed mannosidase amino acid sequences may be employed as immunogens for producing antibodies.

Because naturally occurring epitopes on proteins frequently comprise amino acid residues that are not adjacently arranged in the peptide when the peptide sequence is viewed as a linear molecule, it may be advantageous to utilize longer peptide fragments from the mannosidase amino acid sequences for producing antibodies. Thus, for example, peptides that comprise at least 10, 15, 20, 25, or 30 consecutive amino acid residues of the amino acid sequence may be employed. Monoclonal or polyclonal antibodies to the intact mannosidase, or peptide fragments thereof may be prepared as described below.

#### **B. Monoclonal Antibody Production by Hybridoma Fusion**

Monoclonal antibody to any of various epitopes of the mannosidase enzymes that are identified and isolated as described herein, can be prepared from murine hybridomas according to the classical method of Kohler & Milstein, *Nature*, 256:495, 1975, or a derivative method thereof. Briefly, a mouse is repetitively inoculated with a few micrograms of the selected protein over a period of a few weeks. The mouse is then sacrificed, and the antibody-producing cells of the spleen isolated. The spleen cells are fused by means of polyethylene glycol with mouse myeloma cells, and the excess unfused cells destroyed by growth of the system on selective media comprising aminopterin (HAT media). The successfully fused cells are diluted and aliquots of the dilution placed in wells of a microtiter plate where

growth of the culture is continued. Antibody-producing clones are identified by detection of antibody in the supernatant fluid of the wells by immunoassay procedures, such as ELISA, as originally described by Engvall, *Enzymol.* 70:419, 1980, or a derivative method thereof. Selected positive clones can be expanded and  
5 their monoclonal antibody product harvested for use. Detailed procedures for monoclonal antibody production are described in Harlow & Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory, New York, 1988.

### C. Polyclonal Antibody Production by Immunization

10 Polyclonal antiserum containing antibodies to heterogenous epitopes of a single protein can be prepared by immunizing suitable animals with the expressed protein, which can be unmodified or modified, to enhance immunogenicity. Effective polyclonal antibody production is affected by many factors related both to the antigen and the host species. For example, small molecules tend to be less  
15 immunogenic than other molecules and may require the use of carriers and an adjuvant. Also, host animals vary in response to site of inoculations and dose, with both inadequate or excessive doses of antigen resulting in low-titer antisera. Small doses (ng level) of antigen administered at multiple intradermal sites appear to be most reliable. An effective immunization protocol for rabbits can be found in  
20 Vaitukaitis et al., *J. Clin. Endocrinol. Metab.* 33:988-991, 1971.

Booster injections can be given at regular intervals, and antiserum harvested when the antibody titer thereof, as determined semi-quantitatively, for example, by double immunodiffusion in agar against known concentrations of the antigen, begins to fall. See, for example, Ouchterlony et al., *Handbook of Experimental*  
25 *Immunology*, Wier, D. (ed.), Chapter 19, Blackwell, 1973. A plateau concentration of antibody is usually in the range of 0.1 to 0.2 mg/mL of serum (about 12  $\mu$ M). Affinity of the antisera for the antigen is determined by preparing competitive binding curves, as described, for example, by Fisher, *Manual of Clinical Immunology*, Chapter 42, 1980.

#### D. Antibodies Raised by Injection of cDNA

Antibodies may be raised against the mannosidases of the present invention by subcutaneous injection of a DNA vector that expresses the enzymes in laboratory animals, such as mice. Delivery of the recombinant vector into the animals may be achieved using a hand-held form of the Biolistic system (Sanford et al., *Particulate Sci. Technol.* 5:27-37, 1987, as described by Tang et al., *Nature* (London) 356:153-154, 1992). Expression vectors suitable for this purpose may include those that express the cDNA of the enzyme under the transcriptional control of either the human  $\beta$ -actin promoter or the cytomegalovirus (CMV) promoter. Methods of administering naked DNA to animals in a manner resulting in expression of the DNA in the body of the animal are well known and are described, for example, in U.S. Patent Nos. 5,620,896 ("DNA vaccines against rotavirus infections"); 5,643,578 ("Immunization by inoculation of DNA transcription unit"); and 5,593,972 ("Genetic immunization"), and references cited therein.

#### E. Antibody Fragments

Antibody fragments may be used in place of whole antibodies and may be readily expressed in prokaryotic host cells. Methods of making and using immunologically effective portions of monoclonal antibodies, also referred to as "antibody fragments," are well known and include those described in Better & Horowitz, *Methods Enzymol.*, 178:476-496, 1989; Better et al., in Streilein et al., eds., *Advances in Gene Technology: The Molecular Biology of Immune Disease & the Immune response (ICSU Short Reports)*, 10:105, 1990; Glockshuber et al. *Biochemistry* 29:1362-1367, 1990; and U.S. Patent Nos. 5,648,237 ("Expression of Functional Antibody Fragments"), No. 4,946,778 ("Single Polypeptide Chain Binding Molecules"), and No. 5,455,030 ("Immunotherapy Using Single Chain Polypeptide Binding Molecules"), and references cited therein.

#### IV. Using Mannosidase

##### A. *In vitro*

The mannosidases of the present invention are useful for modifying the glycosylation pattern proteins *in vitro*. In particular, the enzymes of the present invention are useful for modifying target proteins. Target proteins are specific proteins that are the desired product of a process, for example, a glycoprotein that will be used as a therapeutic agent.

The *in vitro* modification of a target protein is accomplished by first obtaining one or more of the mannosidases of the present invention. The mannosidase may be purified to achieve specificity of action. The mannosidase is then placed in contact with the target protein, allowing for the modification of the glycosylation pattern of the target protein in a controlled, *in vitro* environment. *In vitro* methods of modifying the glycosylation pattern of a protein are well known in the art, and examples of such methods are provided in U.S. Patent No. 5,834,251 to Maras, et al., herein incorporated by reference.

##### B. *In vivo*

The glycosylation pattern of a protein, a broad class of proteins, or all glycosylated proteins in a host cell can be modified *in vivo* through the use of molecular biology techniques. The provision of the disclosed nucleic acid sequences allows for both the up-regulation of the disclosed mannosides as well as the down-regulation of the disclosed mannosidases.

Up regulation of the protein can be accomplished by transforming a cell with a vector containing one or more of the disclosed mannosidase genes under the control of a promoter. The promoter can be for example a constitutive promoter, a tissue specific promoter, or an inducible promoter. Thus, the RNA encoding the enzyme(s) is produced and translated into the protein. The increased level of mannosidase in the cell will then alter the glycosylation pattern of proteins produced by the cell.

Conversely, the production of mannosidase by a cell can be down-regulated by transforming the cell with a vector encoding an antisense molecule, or a catalytic

nucleic acid molecule that targets the mannosidase specific nucleic acid sequences. This technique will likely cause the cell to produce less mannosidase and therefore, glycoproteins from the cell will be processed to a lesser extent.

In contrast to the above-described *in vitro* methods, *in vivo* methods can be used to produce multiple different proteins in the same cell, all of which display altered glycosylation patterns.

## V. Examples

### Example 1

10       **Identification of *Aspergillus nidulans*  $\alpha$ -1, 2-mannosidase 1A.** Four previously published Class 1  $\alpha$ -1,2-mannosidase protein sequences were aligned, and conserved blocks were identified for PCR primer design. Selection of sequences for primer design was based on high sequence identity between the four proteins, and low codon redundancy. A codon-preference table for *A. nidulans* was used to decrease the codon redundancy in the primers and to select codons more likely to be found in the genomic DNA. Two primers were selected that contained the least redundancy and the highest possible annealing temperature. Amplification of *A. nidulans* genomic DNA with these two primers yielded a 900-bp product. This PCR product was cloned and the DNA sequence was determined for several representative clones. The DNA sequence of clone pGEM42-9 was used to search the GenBank database for sequence homology (BLAST<sup>TM</sup> search). Several high-scoring matches were found with other Class 1  $\alpha$ -1,2-mannosidases, confirming that a portion of the Class 1  $\alpha$ -1,2-mannosidase from *A. nidulans* had been amplified.

20       The 900-bp PCR fragment was used as a template for the PCR amplification of an  $\alpha$ -1,2-mannosidase-specific radiolabeled probe for library screening. An EMBL-3 library of genomic sequences was screened and several positive plaques were re-screened to isolate a single lambda clone containing the entire  $\alpha$ -1,2-mannosidase gene. Two *Bam*H1 subclones were identified by Southern hybridization, using the same PCR-derived probe, that contained the 5' and 3' ends of the gene. These subclones were sequenced by manual and automated sequencing, using a combination of primer walking with specifically designed primers and restriction enzyme subcloning using universal sequencing primers, yielding 2- to 5-

fold redundancy sequencing of the full-length gene and several thousand base pairs (bp) of flanking sequence (Fig. 1).

## Example 2

5           **Characterization of the  $\alpha$ -1, 2-mannosidase 1A gene.** A BLAST<sup>TM</sup> search of the  $\alpha$ -1,2-mannosidase 1A gene revealed two open reading frames (ORFs) that contained significant homology to the Class 1  $\alpha$ -mannosidases. These ORFs were separated by a region of DNA containing several stop codons, which indicated an intron sequence. This intron sequence was verified by reverse-transcription PCR  
10 (RT-PCR) from total RNA and comparison to the genomic DNA. The amplification product derived from the RNA, using a PCR primer pair that flanked the putative intron, was 50 bp shorter than the similar PCR product derived from genomic DNA. Analysis of the putative intron revealed several typical sequence motifs which are characteristic of eukaryotic intron sequences. The 5'-splice site (5'-GTAAGT-3'  
15 (SEQ ID NO: 7)) fit the consensus sequence for filamentous fungi (5'-GTANGT-3' (SEQ ID NO: 8)), and the 3'-splice site (5'-TAG-3') fit the consensus 5'-YAG-3' (Ballance, *Molecular Industrial Mycology – Systems and Applications for Filamentous Fungi*, Leon and Berka (eds.), Dekker, Inc., 1991; Ballance, *Yeast* 2:229-236, 1986; Gurr et al., *Gene Structure in Eukaryotic Microbes*, Kinghorn  
20 (ed.), IRL Press, 1987). The intron also contained an internal lariat sequence (5'-GCTGAC-3' (SEQ ID NO: 9)), located 15 bp upstream of the 3'-splice site, which fit the consensus 5'-(G/A)CT(G/A)AC-3' (SEQ ID NO: 19).

The deduced amino acid sequence of the *A. nidulans*  $\alpha$ -1,2-mannosidase 1A gene was aligned with other published  $\alpha$ -mannosidases in order to identify potential  
25 introns that do not shift the reading frame and do not contain stop codons but do increase the size of the putative gene product. In addition to the previously confirmed intron, two other regions of the ORF that did not align with other published sequences seemed to be "extra" DNA sequences. These regions were represented by large "gaps" in the multiple sequence alignment for all of the other  
30  $\alpha$ -1,2-mannosidase genes used in the alignment. To determine whether these sequences represented introns, or encoded a polypeptide sequence novel to the *A. nidulans*  $\alpha$ -mannosidase, primer pairs were designed that flanked these regions and

used in RT-PCR. The amplification products derived from the RNA (RT-PCR) were the same size as the amplification products derived from genomic DNA, indicating that there were no introns present in these regions. These regions did not contain consensus splice motifs, though sometimes introns were present with less-conserved splice sequences (Gurr et al., *Gene Structure in Eukaryotic Microbes*, Kinghorn (ed.), IRL Press, 1987).

The position of the translational start codon of the first ORF was determined through examination of the DNA sequence near the stop codon that defined the 5' end of the ORF. A potential in-frame start codon (ATG) occurred 42 bp after the beginning of the ORF, while the next in-frame ATG codon was located 423 bp downstream. The first ATG codon was thus a good candidate for the translational start codon. Translation originating at the first start codon would produce a protein product with an N-terminus larger than other fungal  $\alpha$ -mannosidases, but similar in size to the N-termini of mammalian and insect Class 1  $\alpha$ -mannosidases.

Though no strong consensus sequence surrounds the translational start codon in filamentous fungi (Ballance, *Yeast* 2:229-236, 1986), there is a preference (97%) for a purine at position -3. The putative start codon for this gene has a G at position -3, and thus conforms to this rule. There is a TATA-like element at position -47 and several pyrimidine-rich blocks within 100 bp upstream of the putative start codon. These pyrimidine rich regions are often found in fungal promoters and may influence the level of transcription (Ballance, *Yeast* 2:229-236, 1986). The 5'-untranslated region does not contain a CAAT-box upstream of the TATA-like element, but this is not unusual for fungal promoters.

The  $\alpha$ -1,2-mannosidase 1A protein contained several charged N-terminal amino acids representative of a typical signal sequence motif. A Kyte-Doolittle hydrophathy plot showed that the signal sequence was followed by a highly hydrophobic region approximately 15-16 amino acids in length, likely encoding a transmembrane domain, while the rest of the protein was relatively hydrophilic (Fig. 3). This  $\alpha$ -1,2-mannosidase 1A protein from *A. nidulans* likely forms a type II transmembrane protein, which is a characteristic of other Class 1  $\alpha$ -1,2-mannosidases.



**Example 3**

**Identification of *A. nidulans*  $\alpha$ -1,2-mannosidase 1B.** The deduced coding region of the *A. nidulans*  $\alpha$ -1,2-mannosidase 1A gene was used to search the *A. nidulans* EST Sequencing Project Database using the BLAST™ algorithm to determine if there were multiple  $\beta$ -1,2-mannosidase genes were expressed in this organism. A sequence tag, containing 200 bp of information, was identified that showed significant similarity, but not 100% identity, to the  $\alpha$ -1,2-mannosidase 1A gene, and thus represented a separate gene. This sequence tag was used to design new PCR primers specific for the novel gene ( $\alpha$ -1,2-mannosidase 1B). The primers were designed in regions that, based on multiple sequence alignments of published  $\alpha$ -1,2-mannosidase sequences, would not be expected to be well conserved, and thus would be gene-specific. The 200-bp amplification product from these primers was radiolabeled and used to probe the *A. nidulans* genomic library to select the full-length gene.

A single lambda clone containing the full-length  $\alpha$ -1,2-mannosidase 1B gene was isolated and subcloned into plasmid DNA. A single 5.6-kb *Bam*HI clone containing the full-length gene was identified by Southern analysis. A series of restriction enzyme resections of the 5.6-kb clone were derived and the sequence was determined by automated fluorescence sequencing using the universal priming sites of the cloning vector. The full-length gene and several hundred bases of flanking sequence were sequenced with 2- to 4-fold redundancy at each base position (Fig. 2).

**Example 4**

**Characterization of  $\alpha$ -1,2-mannosidase 1B.** The genomic DNA sequence for the *A. nidulans*  $\alpha$ -1,2-mannosidase 1B gene was searched for open reading frames that showed homology to other  $\alpha$ -1,2-mannosidase genes. The gene contained three open reading frames with homology with Class 1  $\alpha$ -1,2-mannosidases separated by two regions that contained several stop codons and caused a shift in the reading frame containing  $\alpha$ -1,2-mannosidase homology. The two regions that disrupted the open reading frame were analyzed to identify

potential introns. Comparison of the reading frame of the gene with the *A. satoi* and *P. citrinum* genes verified that these regions appeared to contain extra DNA sequences, since the alignment seems to be disrupted near these regions. These regions were searched for consensus intron-splice sites. The first potential intron did indeed contain 5' and 3' consensus splice sites as well as an internal lariat consensus sequence. The sequence 5'-GTACGT-3' (SEQ ID NO: 11) fit the filamentous fungal consensus for a 5'-splice site (5'-GTANGT-3'), and the sequence 5'-TAG-3' fit the consensus for the 3'-splice site (5'-YAG-3'). This intron also contained the internal lariat sequence 5'-ACTGAC-3' (SEQ ID NO: 12) located 11 bp upstream of the 3'-splice site. The second putative intron also contained a consensus 5'-splice site (5'-GTACGT-3'), a consensus 3'-splice site (5'-CAG-3'), and an internal lariat consensus (5'-ACTGAC-3') which was located 11 bp upstream of the 3'-splice site. The two putative introns were verified by RT-PCR. In both cases, the RNA-derived amplification product was smaller than the DNA-derived amplification product, and the size difference corresponded to the predicted size of each intron. Interestingly, the position of each intron in the coding region is correlated directly to two of the introns found in the *P. citrinum*  $\alpha$ -1,2-mannosidase gene, although the intron sequences themselves did not appear to be conserved. The N-terminal region of the first ORF likely contained the translational start site, which would probably be the first methionine after the stop codon defining the beginning of the ORF. Comparison of the coding region of this ORF with the *A. satoi* and *P. citrinum*  $\alpha$ -1,2-mannosidase genes showed that the putative translational start site correlated positionally with the start sites of these genes. This start codon contained a purine at the -3 position, as expected (Ballance, *Yeast* 2:229-236, 1986). A TATA-like element was found at position -76 (TATAT), and the upstream region contained several CT-rich tracts in the sense strand. Seven copies of the sequence CTCC appeared in the 100-bp region upstream of the TATA-like element and could represent an important promoter element. Additionally, there was a 7-bp direct repeat located immediately upstream of the TATA-like element (CCTCAT). The N-terminus of this protein contained a typical signal sequence for insertion into the endoplasmic reticulum and, as seen in a Kyte-Doolittle hydropathy plot (Fig. 3), contained a 10-15 bp hydrophobic region which likely encoded a transmembrane

domain. This protein, like other Class 1  $\alpha$ -1,2-mannosidases, is likely a type-II transmembrane protein normally localized in the ER or Golgi apparatus.

#### Example 5

- 5           **Identification of *A. nidulans*  $\alpha$ -1,2-mannosidase 1C:** A second lambda clone containing the full length  $\alpha$ -1,2-mannosidase 1C gene was recovered and two non-overlapping *Bam*HI subclones (4 kb and 6 kb) were isolated, which together contained the gene and flanking regions. Again, the sequence across the *Bam*HI subcloning junction was verified to eliminate the possibility of missing sequence.
- 10   The gene and several hundred bp of flanking region were fully sequenced (Accession #: AF233287).

#### Example 6

- Comparison of  $\alpha$ 1,2-mannosidase 1A, 1B, and 1C sequences:** The DNA
- 15   sequences of the three Class I  $\alpha$ -1,2-mannosidases were analyzed to determine the amino acid coding sequence, including determination of the correct reading frame, identification of potential intron sequences, and identification of the correct translational start codon for each gene. A BLAST search of the  $\alpha$ -1,2-mannosidase 1A gene revealed two open reading frames (ORFs) separated by a region of DNA
- 20   containing several stop codons which could indicate the presence of an intron sequence. To verify the presence of an intron at this site, PCR products spanning the putative intron were amplified from reverse-transcribed RNA (RT-PCR), cloned into vector and sequenced. Comparison of the  $\alpha$ -mannosidase 1A sequence with RT-PCR sequence verified the presence of a 50 bp intron at the expected splice junction.
- 25   The intron contained a 5'-splice site (5'-GTAAGT-3') which matched the consensus sequence for filamentous fungi (5'-GTANGT-3'), and a 3'-splice site (5'-TAG-3') which matched the consensus 5'-YAG-3' (Ballance, Transformation Systems for Filamentous Fungi and an Overview of Fungal Gene Structure, In Leong, S.A. and Berka, R.M. (eds.), *Molecular Industrial Mycology-Systems and Applications for*
- 30   *Filamentous Fungi*, Dekker, Inc., New York, NY, pp. 1-29, 1991, 1986; Gurr et al., The Structure and Organization of Nuclear Genes of Filamentous Fungi, In Kinghorn (ed.), *Gene Structure in Eukaryotic Microbes*, IRL Press, Oxford, WA, pp.

93-139, 1987). The intron also contained an internal lariat sequence (5'-GCTGAC-3'; SEQ ID NO: 9), located 15 bp upstream of the 3'-splice site, consistent with the consensus 5'-(G/A)CT(G/A)AC-3'(SEQ ID NO: 19) for fungal introns.

The deduced amino acid sequence of the *A. nidulans*  $\alpha$ -1,2-mannosidase 1A gene was aligned with other published  $\alpha$ -mannosidases to determine whether there might have been other introns which did not shift the reading frame and did not contain stop codons but did increase the size of the putative gene product. Two additional regions of the first ORF did not align with other published sequences and appeared as large 'gaps' in the multiple sequence alignment for all of the other  $\alpha$ -1,2-mannosidase genes used in the alignment. To determine whether these sequences represented introns or encoded polypeptide sequence novel to the *A. nidulans*  $\alpha$ -mannosidase, RT-PCR products were compared to genomic PCR products. The PCR amplification products derived from both RNA and genomic DNA, using primer pairs which flanked these regions, were the same size, indicating that there were no introns present in these regions. Although introns can sometimes be present with less conserved splice sequences (Gurr et al., The Structure and Organization of Nuclear Genes of Filamentous Fungi, In Kinghorn (ed.), *Gene Structure in Eukaryotic Microbes*, IRL Press, Oxford, WA, pp. 93-139, 1987), these regions did not contain consensus splice motifs.

The  $\alpha$ -1,2-mannosidase 1B gene contained three open reading frames separated by two regions which contained several stop codons and caused a shift in the reading frame. These regions were searched for consensus intron splice sites. Both of the putative introns contained consensus 5'- and 3'- splice sites and a consensus internal lariat sequence. The two putative introns were verified by RT-PCR. In both cases, the RNA-derived amplification product was smaller than the DNA-derived amplification product, and the size difference corresponded to the predicted size of each intron. Sequencing of the RT-PCR products verified the presence of a 51 bp and 53 bp intron at the respective splice sites. While the position of each intron in the coding region correlated directly to two of the introns found in the *P. citrinum*  $\alpha$ -1,2-mannosidase gene, the intron sequences themselves were not conserved. The  $\alpha$ -1,2-mannosidase 1C gene contained a single contiguous open

reading frame and did not contain any consensus intron sequences. This gene did not appear to contain any intron sequences.

The authentic start codon (ATG) of the  $\alpha$ -1,2-mannosidase 1A gene was inferred by sequence context of the putative start codons, combined with protein  
5 sequence alignments with known  $\alpha$ -mannosidase proteins. A potential start codon (ATG) occurred in frame 42 bp after the stop codon which defined the 5' end of the first ORF, while the next in frame ATG codon was 423 bp downstream. This first codon was thus a better candidate for the translational start codon. Translation originating at this start codon would produce a protein product with an N-terminus  
10 which was larger than other fungal  $\alpha$ -mannosidases, but similar in size to the N-termini of mammalian and insect Class I  $\alpha$ -mannosidases. Although there is not a strong consensus sequence surrounding the translational start codon in filamentous fungi (Ballance, *Yeast* 2:229-236, 1986), there is a preference for a purine at position -3 (97%). The putative start codon for this gene had a G at position -3, and thus  
15 conforms to this rule. Placement of the translational start codon can also be inferred from sequence context with respect to promoter elements and the transcription start site. There was a TATA-like element at position -47 of the putative start site and several pyrimidine rich blocks within 100 bp upstream of the proposed start codon. These pyrimidine rich regions are often found in fungal promoters and may  
20 influence the level of transcription (Ballance, *Yeast* 2:229-236, 1986). The 5' non-translated region did not contain a CAAT-box upstream of the TATA-like element, but this is not unusual for fungal promoters.

The first potential translational start codon in the  $\alpha$ -1,2-mannosidase 1B gene occurred 42 bp into the first ORF. Comparison of the coding region of this ORF  
25 with the *A. satoi* and *P. citrinum*  $\alpha$ -1,2-mannosidase genes showed that the position of the putative translational start site correlated with the start sites of these genes. This start codon also contained a purine at the -3 position, a TATA-like element at position -76, and several CT-rich blocks in the sense strand. The first potential translational start codon of the  $\alpha$ -mannosidase 1C gene occurred 38 bp into the  
30 ORF. The start site also contains a purine at position -3 and CCAAT motif at -221, but did not contain a clearly definable TATA box.

The putative coding region of the  $\alpha$ -1,2-mannosidase 1A gene encodes a 816 amino acid protein with a predicted molecular weight (MW) of 91 kD. This is somewhat larger than other Class I  $\alpha$ -mannosidases, which range in size from 53 kD for the *P. citrinum*  $\alpha$ -mannosidase I (Yoshida et al., *Biochim. Biophys. Acta* 1263:159-162, 1995) to 73 kD for the *H. sapiens*  $\alpha$ -mannosidase 1B (Bause et al., *Eur. J. Biochem.* 217:535-540, 1993), *M. musculus*  $\alpha$ -mannosidase 1A (Lal et al., *J. Biol. Chem.* 269:9872-9881, 1994) and  $\alpha$ -mannosidase 1B (Herscovics et al., *J. Biol. Chem.* 269:9864-9871, 1994), and *S. scrofa*  $\alpha$ -mannosidase I (Bieberich et al., *Eur. J. Biochem.* 246:681-689, 1997). The coding region of the  $\alpha$ -1,2-mannosidase 1B gene encodes a 505 aa protein with a predicted MW of 56 kD, while the  $\alpha$ -1,2-mannosidase 1C gene encodes a 586 aa protein with a predicted MW of 65 kD. Both of these predicted sizes are within the range of currently identified Class I  $\alpha$ -mannosidases.

The putative  $\alpha$ -1,2-mannosidase 1A, 1B, and 1C proteins contained several charged N-terminal amino acids representative of a typical signal sequence motif downstream of the putative start site. Kyte-Doolittle hydropathy plots showed that the signal sequences were followed by highly hydrophobic regions approximately 15-16 amino acids in length, while the remainder of the C-termini were relatively hydrophilic. These proteins likely form type II transmembrane proteins, which is a characteristic of other Class I  $\alpha$ -1,2-mannosidases.

#### Example 7

**Mannosidase activity assays:** Crude protein extracts were obtained from protoplasts and from culture filtrates. Secreted proteins were precipitated from 1.5 mL of culture filtrate by saturation with ammonium sulfate. After centrifugation, the protein extract was resuspended in 500 mL 0.01 M phosphate buffer (pH 6.0) and precipitated again using ammonium sulfate (to ensure removal of sugar residues which would interfere with mannosidase assays). Proteins were resuspended in 100  $\mu$ L 0.01 M phosphate buffer (pH6.0) and 27  $\mu$ L of the extract was used in the mannosidase assay. Intracellular proteins were extracted from protoplasts, which were prepared as in Eades et al., *Glycobiol.* 8:17-33, 1998. Protoplasts were

centrifuged and resuspended in 200  $\mu$ L 0.01 M phosphate buffer (pH6.0)/1% octylthiogluconide, and resuspensions were vortexed vigorously to ensure complete protoplast lysis. Cellular debris was removed by centrifugation, and 27  $\mu$ L of the lysate was used in mannosidase assays.

5 Mannosidase assays were performed using the disaccharide Man- $\alpha$ -1,2-Man- $\alpha$ -OCH<sub>3</sub> as a substrate in a coupled enzyme assay as described earlier (Scaman et al., *Glycobiol.* 6:265-270, 1996), with some modifications. Digestion of the substrate was performed in a 30  $\mu$ L final volume containing 27  $\mu$ L of crude extract in 0.01M phosphate buffer (pH6.0) and 3  $\mu$ L 100 mM disaccharide Man- $\alpha$ -1,2-  
10 Man- $\alpha$ -OCH<sub>3</sub> incubated at 37°C for 3 hours. Detection of released mannose was achieved by addition of 30  $\mu$ L Tris-HCl (pH7.6) and 240  $\mu$ L of developing solution, containing glucose oxidase (55 U/mL), horseradish peroxidase (1 U/mL) and o-dianisidine dihydrochloride (70  $\mu$ g/mL), incubated at 37°C for 3 hours. Absorbance measurements at 450 nm determined final color change. Standard blanks included  
15 all components of the colorimetric reaction, plus the substrate. As a control, enzyme extracts which were not used in the mannosidase digestion were subjected to the colorimetric reaction, to determine the absorbance which is due to the extract itself, and not due to mannose release. These values were subtracted from the absorbance values of the assays. Free mannose was used as a standard. All assays were  
20 performed in triplicate and the mean and standard deviation was calculated for each sample. Mannosidase activity was standardized by comparison with total protein in the crude enzyme extracts, and was defined as the amount of mannose released from substrate per  $\mu$ g of total protein per hour. Protein concentrations were determined by the Bradford method (Bradford, *Anal. Biochem.* 72:248-254, 1976) using BSA as a  
25 standard.

The method described above was used to verify that  $\alpha$ -1,2-mannosidases (1A, 1B, and 1C; SEQ ID NOS: 3, 6, and 18, respectively) were actively expressed intracellularly. These assays were performed to determine the secreted and/or intracellular  $\alpha$ -1,2-mannosidase activity levels. To determine secreted  $\alpha$ -1,2-  
30 mannosidase activity, crude protein extracts were obtained from liquid culture filtrates. Intracellular proteins were extracted from fungal protoplasts using a

detergent buffer. Significant  $\alpha$ -1,2-mannosidase activity was found in the intracellular protein extract (28.16 nmol mannose released from substrate/ $\mu$ g total protein/h; SD = 4.45), whereas very little activity was found in the extracellular extracts (0.77 nmol/ $\mu$ g/h; SD = 0.74). This is consistent with the hypothesis that the

5 Class I  $\alpha$ -1,2-mannosidase genes from *A. nidulans* encode type-II transmembrane proteins which are expressed intracellularly.

### Example 8

**Enzyme and protein assays.** La et al., *Glycobiol.* 8:981-995, 1998, described a protocol that can be used to assay for mannosidase activity. This

10 protocol involved assays for testing the activity of two separate mannosidases.

Briefly, La et al., *Glycobiol.* 8:981-995, 1998, assayed for mannosidase activity using the disaccharide Man- $\alpha$ -1,2-man $\alpha$ -O-CH<sub>3</sub>, as a substrate and the glucoseoxidase/peroxidase coupled enzyme assay as described earlier (Scaman et

15 al., *Glyco. Biol.* 6:265-270, 1996) with some modifications. Assays testing mannosidase activity in crude culture medium, or purified enzyme preparations were performed in flat-bottomed 96-well microtiter plates in a 25- $\mu$ L total reaction volume containing 0.01 M potassium phosphate, pH 6.0, 2-10 mM Man- $\alpha$ -1,2-man- $\alpha$ -O-CH<sub>3</sub> and 5-15  $\mu$ L of enzyme solution at 37°C for 30 minutes, or as specified.

20 The reaction was terminated by adding 25  $\mu$ L of 1.25 M Tris-CL, pH 7.6.

The La et al., *Glycobiol.* 8:981-995, 1998, also provided an additional assay, that was used to test the activity of a second mannosidase. This second mannosidase was assayed in culture medium following 10-fold concentration using a Centricon 30 ultrafiltration membrane (Amicon). Assays were performed in 0.5- $\mu$ L microfuge

25 tubes in 30  $\mu$ L total reaction volume containing 0.01 M potassium phosphate, pH 6.0, 2-10 mM Man $\alpha$ 1,2 Man $\alpha$ -O-CH<sub>3</sub>, and 15-24  $\mu$ L of concentrated culture medium at 37°C for 2 hours, or as specified. The reaction was terminated by heating to 100°C for 5 minutes followed by centrifugation at 16,000 x g for 5 minutes. The supernatants (25  $\mu$ L) were transferred to microtiter plate wells containing 25  $\mu$ L of

30 1.25 M Tris-CL, pH 7.6. In both cases, the amount of mannose released was detected by incubation with 250  $\mu$ L of "developing solution" containing glucose oxidase (55 U/mL), horseradish peroxidase (1 purpurogallin unit/mL), and o-



dianisidine dihydrochloride (70 µg/mL) for 3 hours and 37°C. The final color intensity was determined by measuring the absorbance at 450 nm on a Bio-Tek (Winooski, VT) microtiter plate reader. Free mannose was used as a standard. One unit of mannosidase activity is defined as the amount of enzyme that releases 1 nmol of mannose in 1 minute at 36°C.

### Example 9

**α1,2-Mannosidase activity assays.** Herscovics, and Jelinek-Kelly, *Anal. Biochem.* 166:85-89, 1987, describe an assay that can be used to test mannosidase activity from cell lysates. This assay can be used to test the activity of mannosidases of the present invention.

Briefly, lysates from Sf9 cells (Acc. # CRL-1711) were used for assays of α-1,2-mannosidase activity with [<sup>3</sup>H]Man<sub>9</sub>GlcNAc as described previously (Herscovics, and Jelinek-Kelly, *Anal. Biochem.* 166:85-89, 1987) with some modifications. The cells were seeded into 25-cm<sup>2</sup> tissue culture flasks (Corning Glass Works, Corning, NY) at a density of 3 million cells per flask and infected at a multiplicity of 10 plaque forming units (PFU) per cell with either a recombinant baculovirus encoding the Sf9 α-1,2-mannosidase cDNA under the control of the polyhedrin promoter, or wild-type baculovirus as a control. The cells were harvested by centrifugation at selected times after infection, washed with 100 mM Na<sup>+</sup>-MES (pH 6.0), and lysed using 1 µL per 30,000 cells of the same buffer containing 0.5% Triton X-100 with or without 80 mM EDTA. Five microliters of the lysates were used in each reaction of the α-1,2-mannosidase activity assays. Cell lysates for negative control reactions were carried out in a total volume of 40 µL containing 75 mM Na<sup>+</sup>-MES (pH 6.0), 0.0625% Triton X-100, and 6000 c.p.m. of [<sup>3</sup>H]Man<sub>9</sub>GlcNAc. In addition, some reactions contained CaCl<sub>2</sub>, MgCl<sub>2</sub>, EDTA. The reactions were incubated at 37°C for 2.5 hours, boiled for 2 minutes, and 250 µL of a 5 mM solution of CaCl<sub>2</sub>, MgCl<sub>2</sub>, and MnCl<sub>2</sub>, and 175 µL of a 2.25 µg/µL solution of concanavalin A (Boehringer-Mannheim, Indianapolis, IN) in 3.8 M NaCl were added to each reaction. The solutions were vortexed gently and incubated at room temperature for two minutes. Then, 1 mL of 25% PEG-8000 was added and the solutions were vortexed and incubated for another 5 minutes at room

temperature. The solutions were centrifuged for 2 minutes at 10,000 x g, 1 mL of the supernatant was added to 4 mL of a scintillation cocktail (Ultima Gold; Packard Instrument Company, Meriden, CT), and radioactivity was measured using a liquid scintillation counter (model LS 6000Icl Beckman Instruments, Inc., Fullerton, CA).  
5 (Herscovics and Jelinek-Kelly, *Anal. Biochem.* 166:85-89, 1987).

**Strains, Media and Growth Conditions:** *Aspergillus nidulans* sporecolor mutant SM222 was grown in CYM liquid medium (10 g glucose; 2 g bactopectone; 1.5 g casamino acids; 1 g yeast extract; 10 mL 100X salt solution, 1 mL 1000X trace  
10 elements, 10 mL 100X vitamin solution, and 10 mL 100X adenine solution per litre). Stock solutions (100X salt, 100X vitamin, 1000X trace elements, 100X adenine) were described in Kalsner et al., *Glycoconjugate J.* 12:360-370, 1995. Strains were maintained on CYM agar, and spore suspensions were obtained by washing cultured CYM agar plates with 8 mL 0.001% Tween 80. Mycelia for DNA isolations were  
15 obtained by inoculating 500 mL liquid CYM agar with  $10^8$  spores, and incubating 24 hours at 30°C with constant agitation (200 rpm).

**Oligonucleotide Primer Design:** The forward primer MANFOR2B was designed by reverse translation of the protein sequence GGLGESFYEY (SEQ ID  
20 NO: 13), and the reverse primer MANREV3B was designed from the complement of the reverse translation of the sequence FXLAETLKLYLY (SEQ ID NO: 14). These protein sequences were conserved between the  $\alpha$ -1,2-mannosidase protein sequences of *S. cerevisiae* (Camirand et al., *J. Biol. Chem.*, 266:15120-15127, 1991), *Mus musculus* 1A (Lal et al., *J. Biol. Chem.*, 269:9872-9881, 1994; Acc. #  
25 U04299), *Homo sapiens* 1A (Bause et al., *Eur. J. Biochem.*, 217:535-540, 1993; Acc. # X74837), and rabbit liver (Lal et al., *J. Biol. Chem.*, 269:9872-9881, 1994). A codon usage table compiled for *A. nidulans* was used to aid in nucleotide selection at degenerate sites.

30 **DNA Isolation and PCR Amplification:** Total genomic DNA was extracted from finely ground freeze-dried mycelia of *A. nidulans* strain SM222. Approximately 400 mg of mycelia were vortexed with 2.5 mL of 50 mM EDTA,

0.2% SDS, and centrifuged for 10 minutes. Then, 85  $\mu$ L of 3 M KOAc, 5 M acetic acid were added to the supernatant. Following a 20-minute incubation on ice, the suspension was re-centrifuged and DNA was isopropanol-precipitated from the supernatant. After resuspension in 100  $\mu$ L TE (10 mM Tris pH 7.5, 1 mM EDTA),  
5 the DNA was extracted once with phenol, twice with chloroform/isoamyl alcohol (24:1) and ethanol precipitated. Each PCR reaction consisted of 10-100 ng of genomic DNA, 50 pmol of each primer, 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl, 50 mM KCl, 0.01% gelatin, 0.1% Triton X-100, 200  $\mu$ M each of dATP, dCTP, dTTP, and dGTP, and 2 units *Taq* DNA polymerase (Perkin-Elmer, Norwalk,  
10 Connecticut) in a final volume of 100  $\mu$ L. Amplification was performed in two stages using a Perkin-Elmer thermal cycler. Five cycles at a lower stringency (56°C) were followed by 30 cycles at higher stringency (63°C). The PCR products were eluted from 1% low-melting-point agarose, cloned into T-Vector (Promega, Madison, Wisconsin) using the T-Vector cloning system and sequenced with  
15 Universal Forward and M13 Reverse primers (Amersham International, Buckinghamshire, England).

**Library Construction and Screening:** The cloned PCR products were used as a template to produce a radiolabeled PCR probe for library screening. Briefly, the  
20 PCR reactions were performed as previously described, except that only 100  $\mu$ M dCTP was used and 0.825 uL  $\alpha$ -<sup>32</sup>P-dCTP (5  $\mu$ Ci) was added to the reactions (reduction of the dCTP concentration prior to addition of the radiolabeled nuclide reduces error due to base misincorporation). The radiolabeled PCR products were purified using the Wizard™ PCR Prep Kit (Promega) and scintillation counted to  
25 assess radioactivity.

A genomic library of *A. nidulans* (SM222) sequences was constructed by digesting genomic DNA with *Bam*HI and ligating the resulting fragments into the similarly digested lambda DNA vector EMPL-3. Concatemers of the ligated DNA were packaged using the Gigapack™ II (Stratagene, La Jolla, California) *in vitro*  
30 packaging system. Approximately 10<sup>5</sup> recombinant lambda plaques were immobilised on nylon membranes (Genescreen Plus™, Dupont, Wilmington, Delaware) and hybridized with the radiolabeled PCR product generated from *A.*

*nidulans*. Single hybridizing clones were selected and rescreened. Lambda DNA was isolated with the Wizard™ Lambda DNA Extraction Kit (Promega), digested with restriction enzymes, and subcloned into the pUC18 cloning vector.

5           **DNA Preparation and Sequencing of *A. nidulans*  $\alpha$ -1, 2-mannosidase 1A:**

Subclones, from the *A. nidulans* library, that contained the full-length  $\alpha$ -1,2-mannosidase genes were identified by Southern analysis and sequenced. Initial manual sequencing of the  $\alpha$ -mannosidase 1A gene was performed according to the dideoxynucleotide method using the T7 sequencing kit (Pharmacia). Initial  
10       sequence data were obtained with the universal priming sites, and with specific sequence primers (primer walking). Final sequence data were provided by subcloning the fragments using various restriction enzymes and sequencing with an ABI373 automated fluorescent sequencer (Applied Biosystems, Foster, California). Open reading frames were identified and aligned with known  $\alpha$ -1,2-mannosidase  
15       sequences. The  $\alpha$ -mannosidase 1B gene was sequenced by cloning restriction fragments of the positive lambda subclone into BlueScript II™ cloning vector (Stratagene, La Jolla California) and sequencing on an ABI373 sequencer, using the universal priming sites of the vector.

20           **Identification of Introns:** Total RNA was extracted from fresh mycelia grown in 150 mL liquid media using the guanidine isothiocyanate method (TRIzol, Molecular Research, Cincinnati, Ohio). Approximately 1 gram of fresh mycelia harvested by suction filtration was homogenized with 6 mL of TRIzol reagent, and incubated at 30°C for 10 minutes. To the homogenate, 1.2 mL chloroform was  
25       added, and the mixture was shaken vigorously for 15 seconds, incubated at 30°C for 2 minutes and centrifuged at 3000 rpm. The upper aqueous phase was removed and 3 mL isopropanol were added to precipitate the RNA. After incubation at 30°C for 10 minutes, the RNA was pelleted by centrifugation and washed with 70% ethanol/diethylpyrocarbonate (DEPC). The RNA pellet was resuspended in 30  $\mu$ L  
30       H<sub>2</sub>O/DEPC and 5  $\mu$ L were used for cDNA preparation and PCR amplification.

Reverse-transcriptase PCR was performed using Superscript II™ reverse transcriptase (Gibco BRL) for first-strand cDNA synthesis using an oligo-dT primer

followed by PCR amplification of the cDNA. The oligo-dT primer was annealed to the RNA by adding 1  $\mu$ L of oligo-dT primer (500  $\mu$ g/mL) to 5  $\mu$ L of RNA (approx. 1-3  $\mu$ g) and 6  $\mu$ L of H<sub>2</sub>O/DEPC. The mixture was heated to 70°C for 10 minutes and then quickly chilled on ice. Reverse transcription was achieved by addition of 4  $\mu$ L "First Strand Buffer" (250 mM Tris-HCl pH 8.3, 375 mM KCl, 15 mM MgCl<sub>2</sub>), 2  $\mu$ L 0.1 M dithiothreitol (DTT), 1  $\mu$ L 10 mM dNTPs, and 1  $\mu$ L (200 U) Superscript II™ reverse transcriptase, followed by incubation at 42°C for 50 minutes. After first-strand cDNA synthesis, the RNA was digested by the addition of 1  $\mu$ L (2 U) RNAase H, and incubated at 37°C for 20 minutes. The cDNA was used to amplify regions containing putative introns for comparison with genomic DNA amplification.

**Sequence Comparison:** The Class-1  $\alpha$ -mannosidase protein sequences used in the sequence comparisons were obtained from GenBank as follows:

15 *Saccharomyces cerevisiae* (Camirand et al., *J. Biol. Chem.*, **266**:15120-15127, 1991; Acc. # M63598), *Aspergillus satoii* (Inoue et al., *Biochim. Biophys. Acta*, **1253**:141-145, 1995; Acc. # D49827), *Penicillium citrinum* (Yoshida and Ichishima, *Biochim. Biophys. Acta*, **1263**:159-162, 1995; Acc. # D45839), *Ophiostoma novo-ulmi* (C.J. Eades and W.E. Hintz, unpublished; Acc. # AF129495), *Drosophila melanogaster* (Kerscher et al., *Dev. Biol.*, **168**:613-626, 1995; Acc. # X82640), *Spodoptera frugiperda* (Kawar et al., *Glycobiol.* **7**:433-443, 1997; Acc. # AF005035), *Homo sapiens* 1A (Bause et al., *Eur. J. Biochem.*, **217**:535-540, 1993; Acc. # X74837), *Homo sapiens* 1B (Tremblay et al., *Glycobiol.* **8**:585-595, 1998; Acc. # AF027156), *Mus musculus* 1A (Lal et al., *J. Biol. Chem.* **269**:9872-9881, 1994; Acc. # U04299), *Mus musculus* 1B (Herscovics et al., *J. Biol. Chem.*, **269**:9864-9871, 1994; Acc. # U03457), and *Sus scrofa* (Bieberich et al., *Eur. J. Biochem.*, **246**:681-689, 1997; Acc. # Y12503). These sequences were aligned using the ClustalW algorithm (Thompson et al., *Nuc. Acid. R.* **22**:4673-4680, 1994) included in the DNASTar computer package (DNASTar, Madison Wisconsin). The sequence similarity matrix and dendrogram were also generated with DNASTar.

Having illustrated and described the principles of the invention in multiple embodiments and examples, it should be apparent to those skilled in the art that the

invention can be modified in arrangement and detail without departing from such principles. We claim all modifications coming within the spirit and scope of the following claims.

**We Claim:**

1. A purified protein having mannosidase activity, comprising an amino  
5 acid sequence selected from the group consisting of:
  - (a) an amino acid sequence selected from the group consisting of SEQ  
ID NOS: 3, 6 and 18;
  - (b) an amino acid sequence that differs from those specified in (a) by  
one or more conservative amino acid substitutions; and
  - 10 (c) amino acid sequences having at least 60% sequence identity to the  
sequences specified in (a).
2. A specific binding agent, that binds to the purified protein of claim 1.
- 15 3. An isolated nucleic acid, molecule encoding a protein according to  
claim 1.
4. A recombinant nucleic acid molecule, comprising a promoter  
sequence operably linked to a nucleic acid sequence according to claim 3.
- 20 5. A cell, transformed with a recombinant nucleic acid molecule  
according to claim 4.
6. The transformed cell of claim 6, wherein the cell is selected from the  
25 group consisting of: an insect cell, a yeast cell, an algae cell, a bacterial cell, a  
mammalian cell, and a plant cell.
7. A transgenic fungus, comprising a recombinant nucleic acid  
according to claim 4.

30

8. A method for altering the glycosylation pattern of a macromolecule, comprising contacting the macromolecule with a purified protein according to claim 1.

5 9. The method of claim 8, wherein contacting the macromolecule is performed *in vitro*.

10 10. A method for producing a macromolecule having an altered glycosylation pattern, comprising:  
(a) providing a transformed cell according to claim 4; and  
(b) allowing the transformed cell to produce the macromolecule.

11. An isolated nucleic acid molecule, comprising a sequence selected from the group consisting of:  
15 (a) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;  
(b) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 1;  
(c) at least 30 contiguous nucleotides of the sequence shown in SEQ ID  
20 NO: 1;  
(d) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;  
(e) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 4;  
(f) at least 30 contiguous nucleotides of the sequence shown in SEQ ID  
25 NO: 4;  
(g) at least 15 contiguous nucleotides of the sequence shown in SEQ ID NO: 17;  
(h) at least 20 contiguous nucleotides of the sequence shown in SEQ ID NO: 17; and  
30 (i) at least 30 contiguous nucleotides of the sequence shown in SEQ ID NO: 17.



12. A method for altering the glycosylation pattern of a macromolecule in a sample, comprising:

- (a) adding a purified protein according to claim 1 to the sample;
- (b) incubating the sample with the purified protein; and
- 5 (c) allowing the purified protein to hydrolytically remove at least one terminal mannoside residue from a macromolecule in the sample.

13. A method for isolating a nucleic acid sequence encoding a mannosidase, comprising:

- 10 (a) hybridizing the nucleic acid sequence under high-stringency conditions to at least 50 contiguous nucleotides of a sequence selected from the group consisting of SEQ ID NOS: 1, 4, and 17; and
- (b) identifying the nucleic acid sequence as one that encodes a mannosidase.

15

14. A mannosidase identified by the method of claim 13.

1/22

AGGGCCTTTTCAGGGGCTGGTTGGCCAAACCTGCCG  
TATTTTCGATTCAGGGCGGGCCCTTAATTCTGGGGA  
CCGATATCCCTTGAACGCGGGGGCAATCAGTTCAAC  
CAACCCAGACCGCTGGGCTTGGTATGAGATTGCGAG  
CATCTATGTCGGTGTACCTTCTGAATGACAATGAAT  
GTATTTTACTTCTCGAAAAGAACCCTTGGGCACTGA  
ATTGTGCGGAGAATGATGCCCTGATTATGATACAAC  
TAGTCCGCTCCGTCAAGCCACAAGGGTCTGGGCGAGT  
CCGCTATAAATCAAAATCGCCTGCACGAACAGACGA  
ATAACCAAGAAAACGCCCCGAGCGCGAGCGTTTCTTC  
TTCCTCTAAGCCTTGCAGCTGGCTCTGCGTCTTTGA  
TCAACCCTTTAGCTGAATTTCCCCAGAACTTCAGCC  
CTCTGCATCCTGTCCTTACCGCAACTCGTTAACCTG  
CGCGACCTCGCGCGACCACAGCCTTAGGTTTCGAGA  
TGCCATGAAAATCAGAAATTGAACCCCTTTCCATT  
ACTATCATTCTCTGCATTCTGCGAGTGATCTGTCCT  
TCGACGTTCCCTTCTTTCCAGCGCTGCGGCGCCTTCA  
CTCTCGTTGCCTACGTTTGACCACGGTCCTACCTCT  
CCTACTGCTGATTATTAGGCTCCTCCCTACGCCTCC  
AATACAGGGAAGTCGCCGGCC

M F R A R R S R I S L V  
F A V  
ATGTTTCGTGCACGACGATCTCGCATCTCGCTGGTG  
TTTGCCGTT

I F V L L I F H F S R L  
A V T  
ATATTTGTCCTCCTCATATTCCACTTTAGCCGTCTC  
GCAGTTACG

Figure 1A

2/22

I S L Q S W V P P P P V  
D H H  
ATCAGCCTGCAATCTTGGGTACCTCCGCCGCCCGTC  
GATCACCAT

N P P F P D Q N L K D P  
Y E N  
AATCCCCCTTTCCCCGACCAGAACCTCAAAGATCCA  
TACGAAAAC

D N S A T G S G A P P P  
A L V  
GACAATAGTGCGACCGGCAGTGGGGCTCCTCCGCCT  
GCGTTGGTA

E P E E Y Q R P P L Y T  
D S D  
GAGCCAGAAGAATACCAACGACCACCACTTTACACA  
GATTCAGAT

D S P T P S K E R L D T  
P S N  
GACAGCCCAACTCCGTCAAAGAACGCCTGGACACC  
CCGAGCAAT

V P S Q E P E F D A A R  
L Q T  
GTCCCATCTCAGGAGCCTGAATTTGATGCCGCCAGA  
CTTCAGACG

**Figure 1B**

3/22

G A Q T Q N K H E D D E  
D I V  
GGTGCGCAGACCCAAAATAAACATGAAGATGATGAG  
GATATTGTC

P I S H W K P M P E R H  
P V S  
CCAATTTCTCACTGGAAGCCGATGCCCCGAACGGCAT  
CCAGTCAGT

P E A L I K L P T G Q S  
K E L  
CCGGAGGCTTTGATCAAGCTGCCAACCGGGCAATCA  
AAGGAACTC

P Q L Q A K F K D E S S  
S D K  
CCCCAACTGCAAGCTAAGTTCAAGGACGAGTCGTCC  
TCGGACAAG

M Q R L Q Q L D T I K S  
A F L  
ATGCAGCGGCTGCAACAACCTTGACACTATCAAGTCG  
GCGTTCTTA

H A W N G Y K I S A M G  
H D E  
CATGCGTGGAACGGTTACAAGATCTCTGCCATGGGT  
CATGATGAG

Figure 1C

4/22

V R P L R G G F K D T F  
N G W  
GTTAGACCTCTGCGCGGTGGTTTCAAGGACACATTC  
AATGGCTGG

G A T L V D A L D T L W  
I M D  
GGCGCGACCCTTGTCGACGCCTTGGATACCCTGTGG  
ATCATGGAT

L K E E F S M A V D Y V  
K K I  
CTCAAAGAGGAGTTCTCCATGGCAGTCGACTACGTC  
AAGAAAATC

D F T T S T K K E I P V  
F E T  
GATTTTACCACCAGCACCAAGAAAGAGATTCCGGTC  
TTTGAAACC

T I R Y L G G M L G A Y  
D I S  
ACTATTCGCTACCTAGGCGGGATGCTCGGGGCCTAT  
GATATTTTCG

G H K Y D I L L E K S V  
E L A  
GGACACAAATACGATATACTTTTGGAAAAGTCTGTT  
GAGCTTGCG

**Figure 1D**

5/22

K R S A G P E K G N W H  
V V A  
GATGTCTTGATGGACGCCTTCGACACACCGAACCGG  
ATGCCAACC

T D V L M D A F D T P N  
R M P  
CTCTATTATAAATGGAGCCCAGAGTATGCTTCAGAG  
TTTCGCCGG

T L Y Y K W S P E Y A S  
E F R  
GGGGACTTTAAGGCTGTTCTCGCCGAGCTTGGCTCT  
CTCTCTCTC

R G D F K A V L A E L G  
S L S  
GAGTTCACGCGTTTGGCGCAGTTGACCAAACAGGAC  
AAGTACTAC

L E F T R L A Q L T K Q  
D K Y  
GATGCAATTGCACGAATCACAAATGAGCTCGAAAAG  
TATCAGGAT

Y D A I A R I T N E L E  
K Y Q  
TTGACAAAGCTTCCCGGCTTGTGGCCTCTCAACCTG  
GACGCATCC

Figure 1E

6/22

D L T K L P G L W P L N  
L D A  
GGGTGCAGGCGAGTTCCCGGCGTCTCGCGAGAGCCT  
GCTGCGGCT

S G C R R V P G V S R E  
P A A  
GGGCAGCCAGTCAGATGGTCCTCTGACGAGATCAAC  
TCGACGAGC

A G Q P V R W S S D E I  
N S T  
TCGGTATCGTATCGTACAAGACAAATTCATGAGGGC  
GGAGAGCCT

S S V S Y R T R Q I H E  
G G E  
GTCCGTCATGACAATGATTCGTTTGAAACGGGTTTT  
CCTGTATCA

P V R H D N D S F E T G  
F P V  
GTCGATACTCGGACTCCTCCCCCAAAGCAAGATTGC  
ACCGGAGGC

S V D T R T P P P K Q D  
C T G  
CTCAACGATCAGCTCTCAGGCATTGACAAGTTCGGA  
CTCGGAGCC

Figure 1F

7/22

G L N D Q L S G I D K F  
G L G  
CTTGGTGACTCTACGTACGAGTACTTACCGAAAGAG  
TATATGTTG

A L G D S T Y E Y L P K  
E Y M  
CTCGGCGGTAACAACGACCAGTACCTCAACATGTAT  
CAGAAGGCC

L L G G N N D Q Y L N M  
Y Q K  
ATGGACACAGTGCGAGAATATCTTGTTTATCAGCCA  
ATGCTCAAG

A M D T V R E Y L V Y Q  
P M L  
AATAATCGCGATGTCCGCTTCTTAGCGACAGTTAGT  
ATGACAAAG

K N N R D V R F L A T V  
S M T  
AGCCTTGATGCAAAACCTCCGGGGCGTACCACTTTC  
GCGTACGAA

K S L D A K P P G R T T  
F A Y  
GGCACTCACCTCACCTGTTTTGCTGGTGGTATGCTT  
GCCATTGGC

Figure 1G



8/22

E G T H L T C F A G G M  
L A I  
GCCAAGTTGTTTGGGCTTGATAAGGATCTAAAGCTG  
GGTAGTCAA

G A K L F G L D K D L K  
L G S  
CTGACGGACGGCTGTGTCTGGGCATATGAAGCCACA  
AAGTCCGGA

Q L T D G C V W A Y E A  
T K S  
ATCATGCCGGAAGCATTCCAACCTGGTCCCTTGTAAG  
AAAGGCGAG

G I M P E A F Q L V P C  
K K G  
CCATGCGAATGGGATGAGGACGCATACTACATGGCC  
ATGGATCCT

E P C E W D E D A Y Y M  
A M D  
TATGCCGACAAGCGGCCAATATCACATAACAAACGC  
TCCGCCGGC

P Y A D K R P I S H N A  
E S S  
CCTGAAAAGGGGAATTGGCACGTCGTCGCCACAGCC  
GAATCGTCT

Figure 1H

9/22

S P Q E D K T Q K S T T  
T E G  
TCGCCCCAGGAAGATAAAACACAGAAATCAACCACT  
ACTGAGGGT

R H T G T T T G A G A L  
S H E  
CGACACACCGGTACAACCTACCGGGGCAGGCGCGCTC  
TCGCACGAG

E F V T G K I L N D R L  
P P G  
GAATTCGTCACGGGAAAAATCCTCAACGACCGACTC  
CCGCCGGGC

M T G I S A R Q Y L L R  
P E A  
ATGACAGGGATCTCGGCTCGGCAGTACCTCCTTCGC  
CCGGAGGCG

I E S V F I M F R L T G  
D P S  
ATCGAGTCTGTCTTCATCATGTTCCGCCTCACGGGC  
GATCCTTCC

W R E K G W K M F Q A V  
D K A  
TGGCGCGAAAAGGGTTGGAAGATGTTCCAGGCTGTC  
GACAAAGCC

Figure 11

10/22

T K T E L A N S A I S D  
V T V  
ACGAAGACGGAGCTGGCGAACTCGGCCATTTCCGAC  
GTAACCGTC

D N P R P V D S M E S F  
W L A  
GATAATCCACGCCCGGTGGACAGTATGGAATCATTC  
TGGCTTGCG

E T L K Y F Y L L F S D  
P S L  
GAGACTCTGAAATACTTCTACCTTCTTTTCAGCGAT  
CCAAGCCTG

V S L D E Y V L  
GTGAGCCTTGAGGAATATGTCTTGTAAGTGATGCTT  
GACTTAATC

*intron 1* N T  
E A H  
GACTGCTTGATGCTGACTTTTCCCTTAGGAACACCG  
AGGCTCATC

P F K R P K Y \*\*\*  
CGTTCAAGCGACCCAGGTACTGAAGTACTAATTTAA  
ATGATCTTTTAGCCTGTATCTATACATGGCCGCTCC  
GCTGTAGAAGCATTGATACCATTAAGACAGTATCGC  
TGCATTTCGTGTACCATTGAGCTTCCAGAGGAACCT  
CTTT

Figure 1J

11/22

TACCATGATCGACAGGTTGTTGCCCATATTGGG  
ACTAGATTTGTTTACCCCAGTCGGAATGTGCCT  
AAAGTGGAAGGTATGATGATGCCCAGGATCGCG  
CCCCAGTCATCAACTCCATCATGGGACGGTCCT  
TGATCCTCAAGGCACGAAGTGGAGATCAGGTCC  
GTAGTGCATATGCATGGCCCATCAGCCTGAAGC  
ACTTCCCCAAGCAAAGTCGAGACTCGGACACCG  
ATGATATCCCTGCTGTCCCTGACTGATGCATAG  
TGCATGCCCCCTGCGCTGGCTCCCCTTTTCACTC  
CGCCTGGTCTCCAGTCTCCACTCCTCACCATTG  
ATGTCTGCCCCCGCCCGCCCTCCATCCTCCATC  
ATTCTTATATCTACGGACTCGGTCACTCGTTAT  
CACTAGAGTCCTTGTTTATTCCTAGTGTTTGCA  
TTCTTACGTGTAGTT

M R T L L A L A A F A  
G F A A  
ATGCGTACGCTTCTCGCTCTCGCGGCCTTCGCG  
GGCTTTGCCGCT

A R V P A Y A I T R P  
V M R S  
GCTAGGGTGCCCGCCTACGCCATCACGCGCCCG  
GTGATGCGCAGT

D S R A D A V K E A F  
S H A W  
GATTCTCGCGCCGACGCTGTCAAGGAGGCCTTT  
TCGCATGCCTGG

Figure 2A

12/22

D G Y Y N Y A F P H D  
E L H P  
GACGGTTACTACAACCTACGCTTTTCCTCATGAC  
GAGCTTCACCCG

I S N G Y G D S R N H  
W G A S  
ATTTCGAACGGTTACGGAGACTCGCGAAACCAC  
TGGGGCGCGTCG

A V D A L S T A I M M  
R N A T  
GCCGTCGACGCTCTATCGACGGCCATCATGATG  
CGCAACGCGACC

I V N Q I L D H I A A  
V D Y S  
ATCGTCAACCAGATCCTTGACCATATTGCTGCT  
GTGGACTACTCC

K T N A M V S L F E T  
T I R Y  
AAGACCAACGCCATGGTAAGTTTGTTCGAGACG  
ACCATCCGGTAC

L A G M I S G Y D L L  
K G P A  
CTCGCGGGCATGATTTCTGGATACGACCTGCTC  
AAAGGCCCTGCG

**Figure 2B**

13/22

A G L V D D S R V D V  
L L E Q  
GCGGGGTTGGTGGACGACAGCAGGGTCGACGTG  
CTTCTAGAGCAG

S Q N L A E V L K F A  
F D T P  
TCGCAGAACCTCGCCGAGGTGCTGAAATTCGCG  
TTCGACACTCCT

S G V P Y N M I N I T  
S G G N  
TCTGGTGTGCCGTACAACATGATTAACATTACT  
TCGGGCGGCAAC

D G A T T N G L A V T  
G T L V  
GACGGCGCCACGACCAACGGGCTGGCCGTGACC  
GGTACCTTGGTG

L E W T R L S D L T G  
N D E Y  
CTGGAGTGGACGCGTCTGTCTGGACTTGACTGGG  
AACGACGAGTAT

A R L S Q R A E D Y L  
L H P E  
GCCCGCTTGAGCCAGAGAGCTGAAGACTACCTT  
CTCCACCCGGAG

**Figure 2C**

14/22

P A Q Y E P F P G L I  
G S A V  
CCAGCGCAGTACGAACCGTTCCCTGGATTGATT  
GGAAGCGCAGTC

N I A D G K L A N G H  
I S W N  
AATATTGCCGACGGCAAGCTCGCCAATGGTCAC  
ATCAGCTGGAAT

G G A D S Y Y E Y L I  
K M Y V  
GGTGGCGCAGACTCGTACTACGAGTACCTGATC  
AAGATGTACGTC

Y D P E R F G L Y R D  
R W V A  
TACGATCCCGAACGCTTTGGCCTCTACCGGGAC  
CGCTGGGTCGCA

A A E S S I N H L A S  
H P S T  
GCTGCCGAGTCGAGCATCAACCATCTGGCTTCG  
CACCCGTCCACC

R P D V T F L A T Y N  
E E H Q  
CGCCCAGACGTGACTTTCTTGGCCACTTACAAC  
GAGGAGCATCAG

Figure 2D

15/22

L G L T S Q H L T C F  
D G G S  
CTGGGCCTGACCAGCCAACACCTGACCTGCTTC  
GACGGTGGAAGC

F L L G G T L L D R Q  
D F V D  
TTTCTGCTTGGTGGGACATTGCTGGACCGCCAG  
GACTTTGTCGAC

F G L D L V A G C H E  
T Y N S  
TTCGGCCTTGACCTTGTCGCCGGCTGCCACGAG  
ACGTACAACCTCG

T L T G I G P E Q F S  
W D P N  
ACTCTGACGGGCATCGGCCCTGAGCAATTCAGC  
TGGGACCCTAAC

G V P D S Q K E L F E  
R A G F  
GGTGTGCCCCGACAGCCAGAAGGAGCTGTTCGAG  
CGCGCAGGCTTC

Y I N S G Q Y I L R P  
E V I E  
TACATCAACAGCGGCCAATACATTCTTCGTCCC  
GAAGTCATCGAG

**Figure 2E**



16/22

S F Y Y A W R V T G D  
G T  
AGCTTCTACTATGCATGGCGCGTCACAGGTGAT  
GGAACGGTACGT

*intron 1*  
**TCACTCAGCGCTGCTTCCGTAGGAAGACCATAC**  
**TGACCGCTTTAG**

Y L E W V W N A F T N  
I N K Y  
TACCTCGAATGGGTGTGGAACGCCTTCACCAAC  
ATCAACAAGTAC

C R T A T G F A G L E  
N V N A  
TGCCGCACTGCGACCGGTTTCGCGGGGCTGGAG  
AACGTCAATGCA

A N G G G R I D N Q E  
S F M F  
GCGAACGGCGGAGGCCGGATCGACAACCAGGAG  
AGTTTCATGTTC

A E V L K Y S F L T F  
A P  
GCAGAGGTGCTGAAGTATTCGTTTTTGACTTTT  
GCTCCTGGTACG

**Figure 2F**

17/22

*intron 2*

TTTCCCTAGTTCTGGTTCACCTGTGGAGAATAT  
TACTGACTGCAG

E D D W Q V Q K G S  
G N T F  
CAGAGGACGACTGGCAGGTGCAGAAGGGCAGTG  
GAAATACGTTTG

V Y N T E A H P F K V  
Y T P Q \*  
TTTATAACACCGAGGCGCACCCGTTTAAGGTGT  
ATACGCCTCAGT

\*\*

AGATAGTACATATTGTGCTCTAGCGTACTGCAT  
GCATAACCTAAATGGCATCACCTACTTACTGAC  
TACTCTACTGACAAGCAGTTGCCTTTATATGCT  
CTTAAGGAGCTTGGGCGGTAAATCGTAATTCAC  
TCTAAACTCCGCGTGCTGTTTCGGGGAAGGCTGG  
GC

Figure 2G

18/22

FIGURE 3A

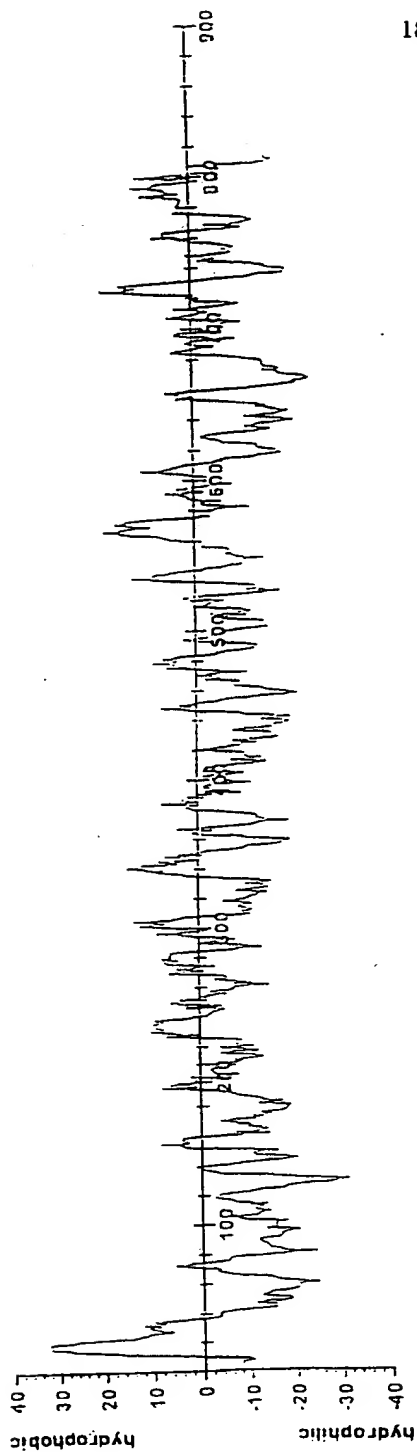
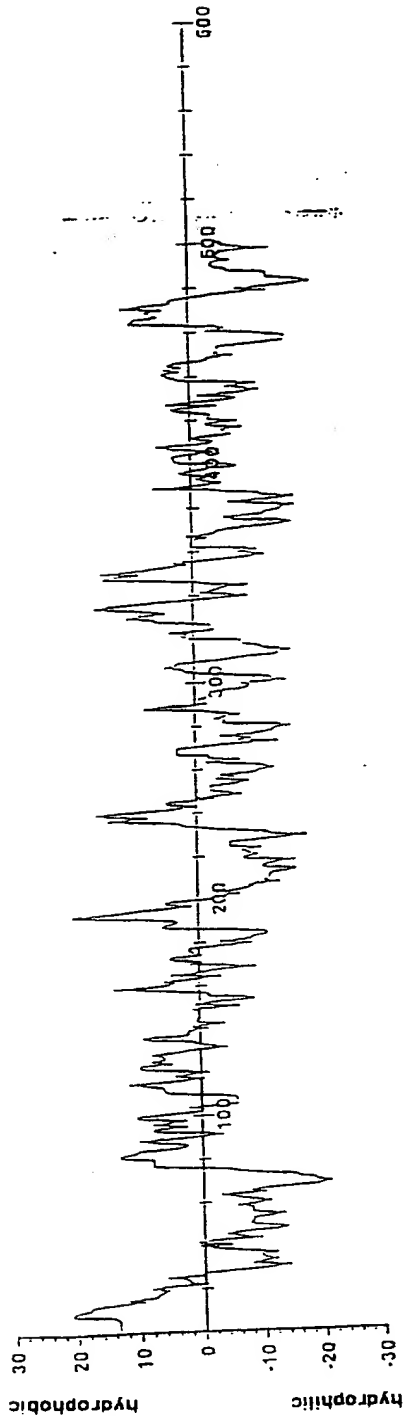


FIGURE 3B



19/22

	1	2	3	4	5	6	7	8	9	10	11	12	13
1- <i>S.cerevisiae</i>		24	24	22	22	21	27	27	28	28	29	28	27
2- <i>A.nidulans</i> 1A			38	28	30	31	23	26	25	25	25	25	24
3- <i>O.novo-ulmi</i> 1A				26	27	26	23	25	22	21	23	21	24
4- <i>A.nidulans</i> 1B					70	66	27	27	26	28	28	28	28
5- <i>A.satoii</i>						72	26	26	29	29	30	30	29
6- <i>P.citrinum</i>							25	26	26	28	28	28	27
7- <i>D.melanogaster</i>								49	42	43	42	43	43
8- <i>S.frugiperda</i>									42	41	43	41	42
9-Human 1A										58	85	57	82
10-Human 1B											60	94	60
11-Mouse 1A												60	85
12-Mouse 1B													58
13- <i>Sacrofa</i>													

FIGURE 4  
SUBSTITUTE SHEET (RULE 26)



21/22

FEATURES	Location/Qualifiers
source	1..2032 /organism="Emericella nidulans" /db_xref="taxon:5072"
gene	36..1796 /gene="MNSIC"
CDS	36..1796 /gene="MNSIC" /note=" [intronless gene]" /codon_start=1 /product="alpha-mannosidase IC"

/translation="MPRRWSSLISITAIFLVLFLLHRNTDTPRAANRATNGPANGF  
ARQOSICSTPPQPPYNRTSTGGFNWGEIPVRYPVSDFIPLSTNSPATLPRIQRSSFPLQSSITKS  
RQAAVKGAQRAWTSYTHAWKADEVRPITAGSRNNFGGWGATLVDNLDTLIMGLD  
EEFAAAVDALADIEFSPHSSPSSSQSTINIFETTIRYLGGLLAAYDLTGCRETRLLD  
KAIQLGEMIIYTSFDTENRMPVPRWNLHKAGNGEPQRAAVQGVLAELASSSLEFTRL  
QLTGDMRYFDAASRITDLLDSQAGHTRIPGLWPVSVNLQKGDLTRGSTFSFGGMADS  
AYEYLGKTYRLLGGVGKGPQYERLARNALDAGIRHLLFRPMPDPHADILLPGVAHAT  
SSSVGLEPRTEHLACFVGGMIALAGKLFNSQTYLDTGRKLTGDCIWYYDNSPLGIMP  
EMFTVPACPSVAECPWDETRGGIYTYVRDGHYFLRPEAMESIFYMWRTGDEKYREA  
AWRMFTAIEAVTKTEFGNAAVRDVMVEEGNVKREDSMESFWMAETLKYLYLIFGETD  
LVSLDDWVFENTEAHPLRGAGS"

BASE COUNT      491 a      554 c      555 g      432 t

# ORIGIN

1 atcgatgtct gctgcataaa ggcagacgga ggaagatgcc gagacggtgg  
tcctccctca  
61 tcagcatcac agccatcttc ttggtcctct tcttcctcct tcataggaat  
acagacacac  
121 cagcgccgc caatagggct acaaacggcc ctgccaacgg  
ctttgctagg cagcaaagca  
181 tatgtccatc aacacccct cagcctccat ataaccgaac  
cagcacgga gggttcaact  
241 ggggtgaaat ccagtcaga taccctgtat ccgacttcat  
cccgtgtca accaactctc  
301 ctgcaacact tccgcgcac caacgtctct ccttcccact  
tcaatcctca atcactaaat  
361 cccgccaggc agcagtcaaa ggtgcctttc agcgcgcacg  
gacctctac acaaccacg  
421 cctggaaggc ggacgagga cggcccatca cggccggatc  
tcgaaacaac tttggcggat  
481 ggggagcgac cctagtcgac aatctcgaca cactgcta  
catggggctg gacgaggagt  
541 tcgcagcggc agtcgacgcg ctcgcagata tagaattcag  
ccgcactcg tccccatcct  
601 cctcccagag cacaatcaac atattcgaaa cgacaatccg  
gtatctgggc ggcttgctcg  
661 cggcgtatga tctactggc tgtcgagaga ctcggctgct  
ggacaaagca atccagcttg

Figure 6A

22/22

721 gggagatgat ctacacctcc ttcgacacag agaaccgcat  
gcccgtacca cgggtggaatc  
781 tgcacaaagc aggcaacgga gagcctcagc gcgcggcagt  
gcagggcgtg ctgctgaac  
841 tcgccagcag cagtctcgag ttcacgcggc tgctgcagct  
gacgggggat atgcggtatt  
901 tcgatgcggc atcccgcatt accgatctgc ttgactccca  
agccggccat acccgatcc  
961 cgggggttggt gccagtcagc gtgaacctgc agaaaggcga  
tctgacctgt gggctgacat  
1021 tcagtttttg cgggatggcc gatagcgcct acgagtatct  
cggcaagacg tatcggtcc  
1081 tcggtggtgt ggggaaagg ccacagtacg agcgtctggc  
gcgaaacgca ctatgccc  
1141 ggattcgaca tctcctcttc cgaccgatga cgcctgatca  
tgcatatc ctctacctg  
1201 gggctgcgca cgcaaccagc tcttcctggt gactcgagcc  
ccggacagag catctgcct  
1261 gttttgtggg tgggatgtac gcgctgcgcg ggaagctttt  
ctcaaaccag acgtacctg  
1321 acaccggccg gaagctgaca gacggttgta tctggtacta  
cgataattca ccgctaggta  
1381 tcatgccgga gatgttcacc gtgccggctt gtccgtcagt  
ggctgaatgt ccttgggacg  
1441 aaacaagggg tggatatctac acctacgtgc gtgatgggca  
ctactttctg cgtcctgagg  
1501 caatggagag tatcttctat atgtggcgca ttacagggga  
cgaaaagtac cgcgaggctg  
1561 catggagaat gttcacggct atcgaagcgg ttacaaagac  
ggagtttggg aatgcggcgg  
1621 tgcgggatgt tatggttgag gaaggaaatg taaagagaga  
agatagcatg gagagtttct  
1681 ggatggcaga gacgttgaag tatctgtatc tgatatttgg  
ggagaccgat ttggtcagct  
1741 tggacgactg ggtgttcaat acggaggcgc accctttgag  
gggtgcaggg agttgacatt  
1801 gtattcacac atcggtatag acaaattata gagtagacgt  
tcaaaacggc caaaactgaa  
1861 tggatagact ccatatgcat tgaatataca atgtattcgc  
tgcaaagcat ggataaaata  
1921 aagatgtaca aagtgtcttt gttgtcgctt tgaaagtggg  
atatcatccc atcataagg  
1981 ggcagtgtaa ccaacctct atataccta catagacagc  
tgatagaccg gc

Figure 6B

## SEQUENCE LISTING

&lt;110&gt; Hintz, William et al.

&lt;120&gt; Mannosidases and Methods for using the Same

&lt;130&gt; 56232

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; 60/157,341

&lt;151&gt; 1999-10-01

&lt;160&gt; 19

&lt;170&gt; PatentIn Ver. 2.0

&lt;210&gt; 1

&lt;211&gt; 3328

&lt;212&gt; DNA

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 1

```

agggcctttt caggggctgg ttggccaaac ctgccgtatt ttcgattcag ggcgggccct 60
taattctggg gaccgatatc ccttgaacgc gggggcaatc agttcaacca acccagaccg 120
ctgggcttgg tatgagattg cgagcatcta tgctgggtga ccttctgaat gacaatgaat 180
gtattttact tctcgaaaag aacccttggg cactgaattg tgcggagaat gatgccctga 240
ttatgataca actagtccgc tccgtcaagc cacaagggtc tgggcagtcc gctataaatc 300
aaaatcgctt gcacgaacag acgaataacc aagaaaacgc ccgagcgcg gctttcttc 360
ttcctctaag ccttgacgtt ggctctgcgt cttgatcaa ccctttagct gaatttcccc 420
agaacttcag ccctctgcat cctgtcctta ccgcaactcg ttaacctgcg cgacctcgcg 480
cgaccacagc cttaggtttc gagatgccat gaaaatcaga aattgaaccc ccttccatt 540
actatcattc tctgcattct gcgagtgatc tgctcctcga cgttccttct ttccagcgct 600
gcggcgccct cactctcggt gcctacgttt gaccacggtc ctacctctcc tactgctgat 660
tattaggtc ctcctacgc ctccaataga gggaagtgc cgccatggt tctgacaga 720
cgatctcgca tctcgctggt gtttgccgtt atattgtcc tctcatatt ccacttagc 780
cgtctcgag ttacgatcag cctgcaatct tgggtacctc cgccgcccg cgtaccat 840
aatcccccct tccccacca gaacctcaa gatccatagc aaaacgaca tagtgcgacc 900
ggcagtgggg ctcctccgcc tgcgttggtg gaggcagaag aataccaacg accaccactt 960
tacacagatt cagatgacag cccaactccg tcaaaagaac gcctggacac cccgagcaat 1020
gtcccattc agagcctga atttgatgcc gccagacttc agacgggtgc gcagaccaa 1080
aataaacatg aagatgatga ggatattgtc ccaatttctc actggaagcc gatgccgaa 1140
cggtatccag tcagtcgga ggctttgatc aagctgcaa ccgggcaatc aaaggaactc 1200
ccccactgc aagctaagtt caaggacgag tctgcctcgg acaagatgca gcggtgcaa 1260
caacttgaca ctatcaagtc ggcgttctta catgcgtgga acggttaca gatctctgcc 1320
atgggtcatg atgaggttag acctctgcgc ggtggttca aggacacatt caatggctgg 1380
ggcgcgaccc ttgtcgacgc cttggatacc ctgtggatca tggatctcaa agaggagttc 1440
tccatggcag tcgactacgt caagaaaatc gattttacca ccagcaccaa gaaagagatt 1500
ccggtctttg aaaccactat tcgtacctta ggcgggatgc tcggggccta tgatatttcg 1560
ggacacaaat acgatatact tttgaaaag tctgttgagc ttgcggatgt cttgatggac 1620
gccttcgaca caccgaaccg gatgccaacc ctctattata aatggagccc agagtatgct 1680
tcagagtttc gccgggggga cttaaggctt gttctcgccg agcttggtc tctctctc 1740
gagttcacgc gtttggcgca gttgacaaa caggacaagt actacgatgc aattgcacga 1800
atcacaaatg agctcgaaa gtatcaggat ttgacaaagc ttcccggtt gtggcctctc 1860
aacctggacg catccgggtg caggcgagtt ccgggctct cgcgagagcc tgcgcggt 1920
gggcagccag tcagatggtc ctctgacgag atcaactcga cgagctcgg atcgtatcgt 1980
acaagacaaa ttcagagggg cggagagcct gtccgtcatg acaatgatc gtttgaaacg 2040
ggttttcctg tatcagtcga tactcgact cctcccccaa agcaagattg caccggagcg 2100
ctcaacgacg agctctcagg cattgacaag ttccgactcg gagcccttgg tgactctacg 2160
tacgagtact taccgaaaga gtatatgttg ctccggcggt acaacgacca gtacctcaac 2220

```



```

atgtatcaga aggccatgga cacagtgcga gaatatcttg ttatcagcc aatgctcaag 2280
aataatcgcg atgtccgctt cttagcgaca gttagtatga caaagagcct tgatgcaaac 2340
cctccggggc gtaccacttt cgcgtacgaa ggcactcacc tcacctgttt tgctgggtgg 2400
atgcttgcca ttggcgccaa gttgtttggg cttgataagg atctaaagct gggtagtcaa 2460
ctgacggacg gctgtgtctg ggcataatgaa gccacaaagt ccggaatcat gccggaagca 2520
ttccaactgg tcccttgtaa gaaaggcgag ccatgcgaat gggatgagga cgcatactac 2580
atggccatgg atccttatgc cgacaagcgg ccaatatcac ataacaaacg ctccgccggc 2640
cctgaaaagg ggaattggca cgtcgtcgcc acagccgaat cgtcttcgcc ccaggaagat 2700
aaaacacaga aatcaaccac tactgagggt cgacacaccg gtacaactac cggggcaggc 2760
gcgctctcgc acgaggaatt cgtcacggga aaaatcctca acgaccgact cccgccgggc 2820
atgacaggga tctcggctcg gcagtacctc cttcgcccgg aggcgatcga gtctgtcttc 2880
atcatgttcc gcctcacggg cgatccttcc tggcgcgaaa aggggttgaa gatgttccag 2940
gctgtcgaca aagccacgaa gacggagctg gcgaactcgg ccatttccga cgtaacccgtc 3000
gataatccac gcccggtgga cagtatgaa tcattctggc ttgaggagac tctgaaatac 3060
ttctaccttc ttttcagcga tccaagcctg gtgagccttg acgaatatgt cttgtaagt 3120
atgcttgact taatcgactg cttgatgctg acttttccct taggaacacc gaggtctatc 3180
cgttcaagcg acccaagtac tgaagtacta atttaaatga tcttttagcc tgatctata 3240
catggccgct ccgctgtaga agcattgata ccattaagac agtatcgctg cattcggtga 3300
ccatttgagc ttccagagga acctcttt 3328

```

&lt;210&gt; 2

&lt;211&gt; 2448

&lt;212&gt; DNA

<213> *Aspergillus nidulans*

&lt;400&gt; 2

```

atgtttcgtg cacgacgac tcgcatctcg ctgggtgttg ccgttatatt tgcctctctc 60
atattccact ttagccgtct cgcagttacg atcagcctgc aatcttggt accctccggc 120
cccgctgac accataatcc ccttttcccc gaccagaacc tcaaagatcc atacgaaaac 180
gacaatagtg cgaccggcag tggggctcct ccgctgcgt tggtagagcc agaagaatac 240
caacgaccac cactttacac agattcagat gacagcccaa ctccgtcaaa agaacgcctg 300
gacaccccca gcaatgtccc atctcaggag cctgaatttg atgccgccag acttcagacg 360
ggtgcgcaga cccaaaataa acatgaagat gatgaggata ttgtccaat ttctcactgg 420
aagccgatgc ccgaacggca tccagtcagt ccggaggctt tgatcaagct gccaacggg 480
caatcaaagg aactccccca actgcaagct aagttcaagg acgagtcgtc ctcggacaag 540
atgcagcggc tgcaacaact tgacactatc aagtcggcgt tcttacatgc gtggaacgg 600
tacaagatct ctgccatggg tcatgatgag gttagacctc tgcgcgggtg tttcaaggac 660
acattcaatg gctggggcgc gacccttgc gacgccttg atacctgtg gatcatggat 720
ctcaaagagg agttctccat ggcagtcgac tacgtcaaga aaatcgattt taccaccagc 780
accaagaaag agattccggg ctttgaacc actattcgct acctaggcgg gatgctcggg 840
gcctatgata tttcgggaca caaatacgat atacttttg aaaagtctgt tgagcttgcg 900
gatgtcttga tggacgcctt cgacacaccg aaccggatgc caacctcta ttataaatg 960
agcccagagt atgcttcaga gtttcgcccg ggggacttta aggctgttct cgcggagctt 1020
ggctctctct ctctcgagtt cagcgcttg gcgcagttga ccaaacagga caagtactac 1080
gatgcaattg ccgaatcac aaatgagctc gaaaagtatc aggatttgac aaagcttccc 1140
ggcttgtggc ctctcaacct ggacgcaccc ggggtgcaggc gagtccccg cgtctcgcga 1200
gagcctgctg cggctgggca gccagtcaga tggctcctcg acgagatcaa ctcgacgagc 1260
tcggtatcgt atcgtacaag acaaattcat gagggcgagg agcctgtccg tcatgacaat 1320
gattcgtttg aaacgggttt tctgtatca gtcgatactc ggactcctcc cccaaagcaa 1380
gattgcaccg gaggctcaa cgatcagctc tcaggcattg acaagttcgg actcggagcc 1440
cttggtgact ctacgtacga gtacttacgg aaagagtata tgttgctcgg cggtacaac 1500
gaccagtacc tcaacatgta tcagaaggcc atggacacag tgcgagaata tcttgtttat 1560
cagccaatgc tcaagaataa tcgcgatgtc cgcttcttag cgacagttag tatgacaaa 1620
agccttgatg cnaaacctcc gggggtacc actttcgcgt acgaaggcac tcacctcacc 1680
tgttttgctg gtggtatgct tgccattggc gccaaagtgt ttgggcttg taagatcta 1740
aagctgggta gtcaactgac ggacggctgt gtctgggcat atgaagccac aaagtccgga 1800
atcatgccgg aagcattcca actggtccct tgtaagaaag gcgagccatg cgaatgggat 1860
gaggacgcat actacatggc catggatcct tatgcccaga agcggccaat atcacataac 1920
aaacgctccg ccggccctga aaaggggaat tggcacgtcg tcgccacagc cgaatcgtct 1980
tcgccccagg aagataaaac acagaaatca accactactg agggtcgaca caccggtaca 2040

```

actaccgggg caggcgcgct ctgcgacgag gaattcgtca cgggaaaaat cctcaacgac 2100  
 cgactcccg cgggcatgac agggatctcg gctcggcagt acctccttcg cccggaggcg 2160  
 atcgagtctg tcttcatcat gttccgcctc acgggcgatc ctctctggcg cgaaaagggt 2220  
 tggaagatgt tccaggctgt cgacaaagcc acgaagacgg agctggcgaa ctcggccatt 2280  
 tccgacgtaa ccgtcgataa tccacgcccc gtggacagta tggaatcatt ctggcttgcg 2340  
 gagactctga aatacttcta ccttcttttc agcgatccaa gcctgggtgag ccttgaggaa 2400  
 tatgtcttga acaccgaggc tcatccgttc aagcgaccca ggtactga 2448

<210> 3

<211> 815

<212> PRT

<213> Aspergillus nidulans

<400> 3

Met Phe Arg Ala Arg Arg Ser Arg Ile Ser Leu Val Phe Ala Val Ile  
 1 5 10 15

Phe Val Leu Leu Ile Phe His Phe Ser Arg Leu Ala Val Thr Ile Ser  
 20 25 30

Leu Gln Ser Trp Val Pro Pro Pro Pro Val Asp His His Asn Pro Pro  
 35 40 45

Phe Pro Asp Gln Asn Leu Lys Asp Pro Tyr Glu Asn Asp Asn Ser Ala  
 50 55 60

Thr Gly Ser Gly Ala Pro Pro Pro Ala Leu Val Glu Pro Glu Glu Tyr  
 65 70 75 80

Gln Arg Pro Pro Leu Tyr Thr Asp Ser Asp Asp Ser Pro Thr Pro Ser  
 85 90 95

Lys Glu Arg Leu Asp Thr Pro Ser Asn Val Pro Ser Gln Glu Pro Glu  
 100 105 110

Phe Asp Ala Ala Arg Leu Gln Thr Gly Ala Gln Thr Gln Asn Lys His  
 115 120 125

Glu Asp Asp Glu Asp Ile Val Pro Ile Ser His Trp Lys Pro Met Pro  
 130 135 140

Glu Arg His Pro Val Ser Pro Glu Ala Leu Ile Lys Leu Pro Thr Gly  
 145 150 155 160

Gln Ser Lys Glu Leu Pro Gln Leu Gln Ala Lys Phe Lys Asp Glu Ser  
 165 170 175

Ser Ser Asp Lys Met Gln Arg Leu Gln Gln Leu Asp Thr Ile Lys Ser  
 180 185 190

Ala Phe Leu His Ala Trp Asn Gly Tyr Lys Ile Ser Ala Met Gly His  
 195 200 205

Asp Glu Val Arg Pro Leu Arg Gly Gly Phe Lys Asp Thr Phe Asn Gly  
 210 215 220

Trp Gly Ala Thr Leu Val Asp Ala Leu Asp Thr Leu Trp Ile Met Asp  
 225 230 235 240

Leu Lys Glu Glu Phe Ser Met Ala Val Asp Tyr Val Lys Lys Ile Asp  
 245 250 255

Phe Thr Thr Ser Thr Lys Lys Glu Ile Pro Val Phe Glu Thr Thr Ile  
 260 265 270  
 Arg Tyr Leu Gly Gly Met Leu Gly Ala Tyr Asp Ile Ser Gly His Lys  
 275 280 285  
 Tyr Asp Ile Leu Leu Glu Lys Ser Val Glu Leu Ala Asp Val Leu Met  
 290 295 300  
 Asp Ala Phe Asp Thr Pro Asn Arg Met Pro Thr Leu Tyr Tyr Lys Trp  
 305 310 315 320  
 Ser Pro Glu Tyr Ala Ser Glu Phe Arg Arg Gly Asp Phe Lys Ala Val  
 325 330 335  
 Leu Ala Glu Leu Gly Ser Leu Ser Leu Glu Phe Thr Arg Leu Ala Gln  
 340 345 350  
 Leu Thr Lys Gln Asp Lys Tyr Tyr Asp Ala Ile Ala Arg Ile Thr Asn  
 355 360 365  
 Glu Leu Glu Lys Tyr Gln Asp Leu Thr Lys Leu Pro Gly Leu Trp Pro  
 370 375 380  
 Leu Asn Leu Asp Ala Ser Gly Cys Arg Arg Val Pro Gly Val Ser Arg  
 385 390 395 400  
 Glu Pro Ala Ala Ala Gly Gln Pro Val Arg Trp Ser Ser Asp Glu Ile  
 405 410 415  
 Asn Ser Thr Ser Ser Val Ser Tyr Arg Thr Arg Gln Ile His Glu Gly  
 420 425 430  
 Gly Glu Pro Val Arg His Asp Asn Asp Ser Phe Glu Thr Gly Phe Pro  
 435 440 445  
 Val Ser Val Asp Thr Arg Thr Pro Pro Pro Lys Gln Asp Cys Thr Gly  
 450 455 460  
 Gly Leu Asn Asp Gln Leu Ser Gly Ile Asp Lys Phe Gly Leu Gly Ala  
 465 470 475 480  
 Leu Gly Asp Ser Thr Tyr Glu Tyr Leu Pro Lys Glu Tyr Met Leu Leu  
 485 490 495  
 Gly Gly Asn Asn Asp Gln Tyr Leu Asn Met Tyr Gln Lys Ala Met Asp  
 500 505 510  
 Thr Val Arg Glu Tyr Leu Val Tyr Gln Pro Met Leu Lys Asn Asn Arg  
 515 520 525  
 Asp Val Arg Phe Leu Ala Thr Val Ser Met Thr Lys Ser Leu Asp Ala  
 530 535 540  
 Asn Pro Pro Gly Arg Thr Phe Ala Tyr Glu Gly Thr His Leu Thr  
 545 550 555 560  
 Cys Phe Ala Gly Gly Met Leu Ala Ile Gly Ala Lys Leu Phe Gly Leu  
 565 570 575

Asp Lys Asp Leu Lys Leu Gly Ser Gln Leu Thr Asp Gly Cys Val Trp  
                   580                                  585                                  590  
 Ala Tyr Glu Ala Thr Lys Ser Gly Ile Met Pro Glu Ala Phe Gln Leu  
                   595                                  600                                  605  
 Val Pro Cys Lys Lys Gly Glu Pro Cys Glu Trp Asp Glu Asp Ala Tyr  
                   610                                  615                                  620  
 Tyr Met Ala Met Asp Pro Tyr Ala Asp Lys Arg Pro Ile Ser His Asn  
                   625                                  630                                  635                                  640  
 Lys Arg Ser Ala Gly Pro Glu Lys Gly Asn Trp His Val Val Ala Thr  
                   645                                  650                                  655  
 Ala Glu Ser Ser Ser Pro Gln Glu Asp Lys Thr Gln Lys Ser Thr Thr  
                   660                                  665                                  670  
 Thr Glu Gly Arg His Thr Gly Thr Thr Thr Gly Ala Gly Ala Leu Ser  
                   675                                  680                                  685  
 His Glu Glu Phe Val Thr Gly Lys Ile Leu Asn Asp Arg Leu Pro Pro  
                   690                                  695                                  700  
 Gly Met Thr Gly Ile Ser Ala Arg Gln Tyr Leu Leu Arg Pro Glu Ala  
                   705                                  710                                  715                                  720  
 Ile Glu Ser Val Phe Ile Met Phe Arg Leu Thr Gly Asp Pro Ser Trp  
                   725                                  730                                  735  
 Arg Glu Lys Gly Trp Lys Met Phe Gln Ala Val Asp Lys Ala Thr Lys  
                   740                                  745                                  750  
 Thr Glu Leu Ala Asn Ser Ala Ile Ser Asp Val Thr Val Asp Asn Pro  
                   755                                  760                                  765  
 Arg Pro Val Asp Ser Met Glu Ser Phe Trp Leu Ala Glu Thr Leu Lys  
                   770                                  775                                  780  
 Tyr Phe Tyr Leu Leu Phe Ser Asp Pro Ser Leu Val Ser Leu Glu Glu  
                   785                                  790                                  795                                  800  
 Tyr Val Leu Asn Thr Glu Ala His Pro Phe Lys Arg Pro Arg Tyr  
                   805                                  810                                  815

&lt;210&gt; 4

&lt;211&gt; 2177

&lt;212&gt; DNA

<213> *Aspergillus nidulans*

&lt;400&gt; 4

cggaatgtgc cttaaagtga aggtatgatg atgcccagga tcgcgccccca gtcatcaact 60  
 ccatcatggg acggtccttg atcctcaagg cacgaagtgg agatcaggtc cgtagtgcac 120  
 atgcatggcc catcagcctg aagcacttcc ccaagcaaag tcgagactcg gacaccgatg 180  
 atatccctgc tgtccctgac tgatgcatag tgcattgccc tgcgctggct ccccttttca 240  
 ctccgcctgg tctccagtct ccactcttca ccattgatgt ctgccccgcg ccgccctcca 300  
 tcttccatca ttcttatatc tacggactcg gtcactcggt atcactagag tccttgttta 360  
 ttcttagtgt ttgcattctt acgtgtagtt atgcgtacgc ttctcgctct cgcggccttc 420  
 gcgggctttg ccgctgctag ggtgcccgcc tacgccatca cgcgcccggg gatgcgcagt 480  
 gattctcgcg ccgacgctgt caaggaggcc ttttcgcatg cctgggacgg ttactacaac 540

```

tacgcttttc ctcacgacga gcttcacccg atttcgaacg gttacggaga ctgcgaaac 600
cactggggcg cgtcgcccg cgcagctcta tcgacggcca tcatgatgag caacgcgacc 660
atcgtaaac agatccttga ccatattgct gctgtggact actccaagac caacgccatg 720
gtaagtgttgc tcgagacgac catccgggtac ctccgcccga tgatttctgg atacgacctg 780
ctcaaagggc ctgcggcggg gttggtggac gacagcaggg tcgacgtgct tctagagcag 840
tcgcagaacc tcgccgaggt gctgaaatc gcgttcgaca ctcttctgg tgtgccgtac 900
aacatgatta acattacttc gggcggaac gacggcgcca cgaccaacgg gctggccgtg 960
accggtacct tgggtgctga gtggacgct ctgtcggact tgactgggaa cgacgagtat 1020
gcccgcctta gccagagagc tgaagactac cttctccacc cggagccagc gcagtacgaa 1080
ccgttccttg gattgattgg aagcgcagtc aatattgccg acggcaagct cgccaatggt 1140
cacatcagct ggaatggtgg cgcagactcg tactacgagt acctgatcaa gatgtacgtc 1200
tacgatcccg aacgcttttg cctctaccgg gaccgctggg tcgcagctgc cgagtcgagc 1260
atcaaccatc tggcttcgca cccgtccacc cggccagacg tgactttctt ggccacttac 1320
aacgaggagc atcagctggg cctgaccagc caacacctga cctgcttcga cgggtggaagc 1380
tttctgcttg gtgggacatt gctggaccgc caggactttg tcgacttcgg ccttgacctt 1440
gtcgcggcgt gccacgagac gtacaactcg actctgacgg gcatcggccc tgagcaattc 1500
agctgggacc ctaacggtgt gccgcagac cagaaggagc tgttcgagcg cgcaggcttc 1560
tacatcaaca gcggccaata cattcttcgt cccgaagtca tcgagagctt ctactatgca 1620
tgccgcgtca cagggtgatg aacggtacgt tcaactcagc ctgcttcctg aggaagacca 1680
tactgaccgc tttagtacct cgaatgggtg tggaaacgct tcaccaacat caacaagtac 1740
tgccgcactg cgaccggttt cgcggggctg gagaacgtca atgcagcgaa cggcggaggc 1800
cggatcgaca accaggagag tttcatgttc gcagaggtgc tgaagtattc gtttttact 1860
tttgcctctg gtacgttttc ctagtcttg ttcacctgtg gagaatatta ctgactgcag 1920
cagaggacga ctggcaggtg cagaaggcga gtggaaatac gtttgattat aacaccgagg 1980
cgcaaccgtt taaggtgtat acgcctcagt agatagtaca tattgtgctc tagcgtactg 2040
catgcataac ctaaatggca tcacctactt actgactact ctactgacaa gcagttgcct 2100
ttatatgctc ttaaggagct tggcgggtaa atcgtaattc actctaaact ccgctgtgctg 2160
ttcggggaag gctgggc

```

&lt;210&gt; 5

&lt;211&gt; 1515

&lt;212&gt; DNA

<213> *Aspergillus nidulans*

&lt;400&gt; 5

```

atgcgtacgc ttctcgtct cgcggccttc gcgggctttg ccgctgctag ggtgcccgcc 60
tacgccatca cgcgcccggt gatgcgcagt gattctcgcg ccgacgctgt caaggaggcc 120
ttttcgcatt cctgggacgg ttactacaac tacgcttttc ctcacgacga gcttcacccg 180
atttcgaacg gttacggaga ctgcgaaac cactggggcg cgtcgcccg cgcagctcta 240
tcgacggcca tcatgatgag caacgcgacc atcgtaaac agatccttga ccatattgct 300
gctgtggact actccaagac caacgccatg gtaagtgttgc tcgagacgac catccgggtac 360
ctcgcgggca tgatttctgg atacgacctg ctcaaaggcc ctgcggcggg gttggtggac 420
gacagcaggg tcgacgtgct tctagagcag tcgcagaacc tcgccgaggt gctgaaatc 480
gcgttcgaca ctcttctgg tgtgccgtac aacatgatta acattacttc gggcggaac 540
gacggcgcca cgaccaacgg gctggccgtg accggtacct tgggtgctga gtggacgctg 600
ctgtcggact tgactgggaa cgacgagtat gccgcttga gccagagagc tgaagactac 660
cttctccacc cggagccagc gcagtacgaa ccgttccttg gattgattgg aagcgcagtc 720
aatattgccg acggcaagct cgccaatggt cacatcagct ggaatggtgg cgcagactcg 780
tactacgagt acctgatcaa gatgtacgtc tacgatcccg aacgcttttg cctctaccgg 840
gaccgctggg tcgcagctgc cgagtcgagc atcaaccatc tggcttcgca cccgtccacc 900
cgcccagacg tgactttctt ggccacttac aacgaggagc atcagctggg cctgaccagc 960
caacacctga cctgctcga cgggtggaagc tttctgcttg gtgggacatt gctggaccgc 1020
caggactttg tcgacttcgg ccttgacctt gtcgccggct gccacgagac gtacaactcg 1080
actctgacgg gcatcggccc tgagcaattc agctgggacc ctaacggtgt gccgcagac 1140
cagaaggagc tgttcgagcg cgcaggcttc tacatcaaca gcggccaata cattcttcgt 1200
cccgaagtca tcgagagctt ctactatgca tggcgctga cagggtatg aacgtacctc 1260
gaatgggtgt ggaacgcctt caccaacatc aacaagtact gccgcactgc gaccggtttc 1320
gcggggctgg agaacgtcaa tgcagcgaac ggcggaggcc ggatcgacaa ccaggagagt 1380
ttcatgtctg cagaggtgct gaagtattcg tttttgactt ttgctcctga ggacgactgg 1440
cagggtgcaga agggcagtg aaatacgttt gtttataaca ccgaggcgca cccgtttaag 1500
gtgtatacgc ctacg

```

&lt;210&gt; 6

&lt;211&gt; 505

&lt;212&gt; PRT

<213> *Aspergillus nidulans*

&lt;400&gt; 6

Met Arg Thr Leu Leu Ala Leu Ala Ala Phe Ala Gly Phe Ala Ala Ala  
 1 5 10 15

Arg Val Pro Ala Tyr Ala Ile Thr Arg Pro Val Met Arg Ser Asp Ser  
 20 25 30

Arg Ala Asp Ala Val Lys Glu Ala Phe Ser His Ala Trp Asp Gly Tyr  
 35 40 45

Tyr Asn Tyr Ala Phe Pro His Asp Glu Leu His Pro Ile Ser Asn Gly  
 50 55 60

Tyr Gly Asp Ser Arg Asn His Trp Gly Ala Ser Ala Val Asp Ala Leu  
 65 70 75 80

Ser Thr Ala Ile Met Met Arg Asn Ala Thr Ile Val Asn Gln Ile Leu  
 85 90 95

Asp His Ile Ala Ala Val Asp Tyr Ser Lys Thr Asn Ala Met Val Ser  
 100 105 110

Leu Phe Glu Thr Thr Ile Arg Tyr Leu Ala Gly Met Ile Ser Gly Tyr  
 115 120 125

Asp Leu Leu Lys Gly Pro Ala Ala Gly Leu Val Asp Asp Ser Arg Val  
 130 135 140

Asp Val Leu Leu Glu Gln Ser Gln Asn Leu Ala Glu Val Leu Lys Phe  
 145 150 155 160

Ala Phe Asp Thr Pro Ser Gly Val Pro Tyr Asn Met Ile Asn Ile Thr  
 165 170 175

Ser Gly Gly Asn Asp Gly Ala Thr Thr Asn Gly Leu Ala Val Thr Gly  
 180 185 190

Thr Leu Val Leu Glu Trp Thr Arg Leu Ser Asp Leu Thr Gly Asn Asp  
 195 200 205

Glu Tyr Ala Arg Leu Ser Gln Arg Ala Glu Asp Tyr Leu Leu His Pro  
 210 215 220

Glu Pro Ala Gln Tyr Glu Pro Phe Pro Gly Leu Ile Gly Ser Ala Val  
 225 230 235 240

Asn Ile Ala Asp Gly Lys Leu Ala Asn Gly His Ile Ser Trp Asn Gly  
 245 250 255

Gly Ala Asp Ser Tyr Tyr Glu Tyr Leu Ile Lys Met Tyr Val Tyr Asp  
 260 265 270

Pro Glu Arg Phe Gly Leu Tyr Arg Asp Arg Trp Val Ala Ala Ala Glu  
 275 280 285

Ser Ser Ile Asn His Leu Ala Ser His Pro Ser Thr Arg Pro Asp Val  
 290 295 300  
 Thr Phe Leu Ala Thr Tyr Asn Glu Glu His Gln Leu Gly Leu Thr Ser  
 305 310 315 320  
 Gln His Leu Thr Cys Phe Asp Gly Gly Ser Phe Leu Leu Gly Gly Thr  
 325 330 335  
 Leu Leu Asp Arg Gln Asp Phe Val Asp Phe Gly Leu Asp Leu Val Ala  
 340 345 350  
 Gly Cys His Glu Thr Tyr Asn Ser Thr Leu Thr Gly Ile Gly Pro Glu  
 355 360 365  
 Gln Phe Ser Trp Asp Pro Asn Gly Val Pro Asp Ser Gln Lys Glu Leu  
 370 375 380  
 Phe Glu Arg Ala Gly Phe Tyr Ile Asn Ser Gly Gln Tyr Ile Leu Arg  
 385 390 395 400  
 Pro Glu Val Ile Glu Ser Phe Tyr Tyr Ala Trp Arg Val Thr Gly Asp  
 405 410 415  
 Gly Thr Tyr Leu Glu Trp Val Trp Asn Ala Phe Thr Asn Ile Asn Lys  
 420 425 430  
 Tyr Cys Arg Thr Ala Thr Gly Phe Ala Gly Leu Glu Asn Val Asn Ala  
 435 440 445  
 Ala Asn Gly Gly Gly Arg Ile Asp Asn Gln Glu Ser Phe Met Phe Ala  
 450 455 460  
 Glu Val Leu Lys Tyr Ser Phe Leu Thr Phe Ala Pro Glu Asp Asp Trp  
 465 470 475 480  
 Gln Val Gln Lys Gly Ser Gly Asn Thr Phe Val Tyr Asn Thr Glu Ala  
 485 490 495  
 His Pro Phe Lys Val Tyr Thr Pro Gln  
 500 505

<210> 7  
 <211> 6  
 <212> DNA  
 <213> Aspergillus nidulans

<400> 7  
 gtaagt

6

<210> 8  
 <211> 6  
 <212> DNA  
 <213> Aspergillus nidulans

<400> 8  
 gtangt

6

<210> 9  
 <211> 6

&lt;212&gt; DNA

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 9

gctgac

6

&lt;210&gt; 10

&lt;211&gt; 6

&lt;212&gt; DNA

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 10

rctrac

6

&lt;210&gt; 11

&lt;211&gt; 6

&lt;212&gt; DNA

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 11

gtacgt

6

&lt;210&gt; 12

&lt;211&gt; 6

&lt;212&gt; DNA

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 12

actgac

6

&lt;210&gt; 13

&lt;211&gt; 10

&lt;212&gt; PRT

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 13

Gly Gly Leu Gly Glu Ser Phe Tyr Glu Tyr  
1 5 10

&lt;210&gt; 14

&lt;211&gt; 9

&lt;212&gt; PRT

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 14

Leu Ala Glu Thr Leu Lys Tyr Leu Tyr  
1 5

&lt;210&gt; 15

&lt;211&gt; 29

&lt;212&gt; DNA

&lt;213&gt; PCR primer

<220> y represents c, t, or u; r represents g or a; n represents a, c,  
g, t, or u; and h represents a, c, t, or u

&lt;400&gt; 15

gggggctng gygartcntt ctacgagta

29



<210> 16  
 <211> 32  
 <212> DNA  
 <213> PCR primer

<220> y represents c, t, or u; r represents g or a; n represents a, c, g, t, or u; and h represents a, c, t, or u

<400> 16  
 gtanaggtac ttnagngtct cngcnagrha gaa

33

<210> 17  
 <211> 2032  
 <212> DNA  
 <213> *Aspergillus nidulans*

<400> 17  
 atcgatgtct gctgcataaa ggcagacgga ggaagatgcc gagacggtgg tcctccctca 60  
 tcagcatcac agccatcttc ttggtcctct tcttcctcct tcataggaat acagacacac 120  
 cacgcgccgc caatagggct acaaacggcc ctgccaacgg ctttgctagg cagcaaaagca 180  
 tatgtccatc aacaccccct cagcctccat ataaccgaac cagcacggga gggttcaact 240  
 ggggtgaaat cccagtcaga taccctgtat ccgacttcat cccgctgtca accaactctc 300  
 ctgcaacact tccgcgcac caacgctctt ccttcccact tcaatcctca atcactaaat 360  
 cccgccaggc agcagtcaaa ggtgcctttc agcgcgcgat gacctcctac acaacccacg 420  
 cctggaaggc ggacgaggta cggcccatca cggccggatc tcgaaacaac ttggcggtat 480  
 ggggagcgac cctagtcgac aatctcgaca cactgctaata catggggctg gacgaggagt 540  
 tcgcagcggc agtcgacgcg ctgcgagata tagaattcag cccgcactcg tccccatcct 600  
 cctccagag cacaatcaac atattcgaaa cgacaatccg gtatctgggc ggcttgctcg 660  
 cggcgatga tctcactggc tgcgagaga ctgcgctgct ggacaaagca atccagcttg 720  
 gggagatgat ctacacctcc ttgcacacag agaaccgcat gcccgtacca cggtggaatc 780  
 tgcacaaagc aggcaacgga gagcctcagc gcgcggcagt gcaggggctg ctgcctgaac 840  
 tcgcagcgac cagctctcgag ttcacgcggc tgcgcgagct gacgggggat atgcgggtat 900  
 tcgatgcggc atcccgcat accgatctgc ttgactccca agccggccat acccggtacc 960  
 cgggggtgtg gccagtcagc gtgaacctgc agaaaggcga tctgacctgt gggtcgacat 1020  
 tcagtttttg cgggatggcc gatagcgccct acgagtatct cggcaagacg tatcggtcc 1080  
 tcggtggtgt ggggaaaggg ccacagtacg agcgtctggc gcgaaacgca ctagatgccg 1140  
 ggattcgaca tctcctcttc cgaccgatga cgcttgatca tgcagatct ctcctaccg 1200  
 gggtcgcgca cgcacaccagc tcttcggtgg gactcgagcc cgggacagag catctcgct 1260  
 gttttgtggg tgggatgtac gcgctcgccg ggaagctttt ctcaaaccag acgtacctcg 1320  
 acaccggccg gaagctgaca gacggttgta tctggtacta cgataattca ccgctaggt 1380  
 tcatgccgga gatgttcacc gtgccggcct gtccgtcagt ggctgaatgt ccttgggacg 1440  
 aaacaagggg tggatatctac acctacgtgc gtgatgggca ctactttctg cgtcctgagg 1500  
 caatggagag tatcttctat atgtggcgca ttacagggga cgaaaagtac cgcgaggctg 1560  
 catggagaat gttcacggct atcgaaagcg ttacaaagac ggagtttggg aatgcggcgg 1620  
 tgcgggatgt tatggttgag gaaggaaatg taaagagaga agatagcatg gagagtttct 1680  
 ggatggcaga gacgttgaaag tatctgtatc tgatatttgg ggagaccgat ttggtcagct 1740  
 tggacgactg ggtgttcaat acggaggcgc accctttgag ggggtgcagg agttgacatt 1800  
 gtattcacac atcggtatag acaaattata gtagtagcgt tcaaaacggc caaaactgaa 1860  
 tggatagact ccataatgcat tgaatataca atgtattcgc tgcaaaagcat ggataaaata 1920  
 aagatgtaca aagtgtcttt gttgtcgctt tgaaagtggg atatcatccc atcataaggt 1980  
 ggcagtgtaa ccaaccctct atatcaccta catagacagc tgatagaccg gc 2032

<210> 18  
 <211> 586  
 <212> PRT  
 <213> *Aspergillus nidulans*

<400> 18

Met Pro Arg Arg Trp Ser Ser Leu Ile Ser Ile Thr Ala Ile Phe Leu  
 1 5 10 15  
 Val Leu Phe Phe Leu Leu His Arg Asn Thr Asp Thr Pro Arg Ala Ala  
 20 25 30  
 Asn Arg Ala Thr Asn Gly Pro Ala Asn Gly Phe Ala Arg Gln Gln Ser  
 35 40 45  
 Ile Cys Pro Ser Thr Pro Pro Gln Pro Pro Tyr Asn Arg Thr Ser Thr  
 50 55 60  
 Gly Gly Phe Asn Trp Gly Glu Ile Pro Val Arg Tyr Pro Val Ser Asp  
 65 70 75 80  
 Phe Ile Pro Leu Ser Thr Asn Ser Pro Ala Thr Leu Pro Arg Ile Gln  
 85 90 95  
 Arg Ser Ser Phe Pro Leu Gln Ser Ser Ile Thr Lys Ser Arg Gln Ala  
 100 105 110  
 Ala Val Lys Gly Ala Phe Gln Arg Ala Trp Thr Ser Tyr Thr Thr His  
 115 120 125  
 Ala Trp Lys Ala Asp Glu Val Arg Pro Ile Thr Ala Gly Ser Arg Asn  
 130 135 140  
 Asn Phe Gly Gly Trp Gly Ala Thr Leu Val Asp Asn Leu Asp Thr Leu  
 145 150 155 160  
 Leu Ile Met Gly Leu Asp Glu Glu Phe Ala Ala Ala Val Asp Ala Leu  
 165 170 175  
 Ala Asp Ile Glu Phe Ser Pro His Ser Ser Pro Ser Ser Ser Gln Ser  
 180 185 190  
 Thr Ile Asn Ile Phe Glu Thr Thr Ile Arg Tyr Leu Gly Gly Leu Leu  
 195 200 205  
 Ala Ala Tyr Asp Leu Thr Gly Cys Arg Glu Thr Arg Leu Leu Asp Lys  
 210 215 220  
 Ala Ile Gln Leu Gly Glu Met Ile Tyr Thr Ser Phe Asp Thr Glu Asn  
 225 230 235 240  
 Arg Met Pro Val Pro Arg Trp Asn Leu His Lys Ala Gly Asn Gly Glu  
 245 250 255  
 Pro Gln Arg Ala Ala Val Gln Gly Val Leu Ala Glu Leu Ala Ser Ser  
 260 265 270  
 Ser Leu Glu Phe Thr Arg Leu Ser Gln Leu Thr Gly Asp Met Arg Tyr  
 275 280 285  
 Phe Asp Ala Ala Ser Arg Ile Thr Asp Leu Leu Asp Ser Gln Ala Gly  
 290 295 300  
 His Thr Arg Ile Pro Gly Leu Trp Pro Val Ser Val Asn Leu Gln Lys  
 305 310 315 320  
 Gly Asp Leu Thr Arg Gly Ser Thr Phe Ser Phe Gly Gly Met Ala Asp

325	330	335
Ser Ala Tyr Glu Tyr Leu Gly Lys Thr Tyr Arg Leu Leu Gly Gly Val		
340	345	350
Gly Lys Gly Pro Gln Tyr Glu Arg Leu Ala Arg Asn Ala Leu Asp Ala		
355	360	365
Gly Ile Arg His Leu Leu Phe Arg Pro Met Thr Pro Asp His Ala Asp		
370	375	380
Ile Leu Leu Pro Gly Val Ala His Ala Thr Ser Ser Ser Val Gly Leu		
385	390	400
Glu Pro Arg Thr Glu His Leu Ala Cys Phe Val Gly Gly Met Tyr Ala		
405	410	415
Leu Ala Gly Lys Leu Phe Ser Asn Gln Thr Tyr Leu Asp Thr Gly Arg		
420	425	430
Lys Leu Thr Asp Gly Cys Ile Trp Tyr Tyr Asp Asn Ser Pro Leu Gly		
435	440	445
Ile Met Pro Glu Met Phe Thr Val Pro Ala Cys Pro Ser Val Ala Glu		
450	455	460
Cys Pro Trp Asp Glu Thr Arg Gly Gly Ile Tyr Thr Tyr Val Arg Asp		
465	470	480
Gly His Tyr Phe Leu Arg Pro Glu Ala Met Glu Ser Ile Phe Tyr Met		
485	490	495
Trp Arg Ile Thr Gly Asp Glu Lys Tyr Arg Glu Ala Ala Trp Arg Met		
500	505	510
Phe Thr Ala Ile Glu Ala Val Thr Lys Thr Glu Phe Gly Asn Ala Ala		
515	520	525
Val Arg Asp Val Met Val Glu Glu Gly Asn Val Lys Arg Glu Asp Ser		
530	535	540
Met Glu Ser Phe Trp Met Ala Glu Thr Leu Lys Tyr Leu Tyr Leu Ile		
545	550	555
Phe Gly Glu Thr Asp Leu Val Ser Leu Asp Asp Trp Val Phe Asn Thr		
565	570	575
Glu Ala His Pro Leu Arg Gly Ala Gly Ser		
580	585	

&lt;210&gt; 19

&lt;211&gt; 6

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;221&gt; variation

&lt;222&gt; (1)

&lt;223&gt; N = G OR A

<220>

<221> variation

<222> (4)

<223> N = G OR A

<220>

<223> Description of Artificial Sequence: CONSENSUS  
SPLICE SITE

<400> 19

nctnac

6

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/27210

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12N 9/00, 9/24, 15/09, 1/20; C07K 16/00  
US CL : 435/69.2, 183, 200; 530/389.1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.2, 183, 200; 530/389.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
BIOSIS, CAPLUS, EMBASE, MEDLINE, SCISEARCH, USPTO WEST.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	HERSCOVIC ET AL. Isolation of a mouse golgi mannosidase cDNA, a member of a gene family conserved from yeasts to mammals. J. Biol. Chem. 01 April 1994, Vol. 269, No. 13, pages 9864-9871, see especially pages 9870-9871.	1, 3-14 ----- 1, 3-14
X --- Y	LAL ET AL. Isolation and expression of murine and rabbit cDNAs encoding an alpha-mannosidase involved in the processing of asparagine-linked oligosaccharides. J. Biol. Chem. 01 April 1994, Vol. 269, No. 13, pages 9872-9881, see entire article.	1, 3-14 ----- 1, 3-14
Y	TURCO ET AL. Altered G-protein glycosylation in vesicular stomatitis virus-infected glucose-deprived baby hamster kidney cells. J. Biol. Chem. 10 August 1982, Vol. 257, No. 15, pages 8674-8679, see entire article.	1, 3-14

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

15 JANUARY 2001

Date of mailing of the international search report

28 FEB 2001

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

MANJUNATH RAO

Telephone No.

(703) 308-0150

TERRY J. DEY  
PARALEGAL SPECIALIST  
TECHNOLOGY CENTER 1600

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/27210

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	US 5,837,836 A (FRIDERICI ET AL.) 17 November 1998 (17/11/98) see entire document.	2 --- 2
X --- Y	US 5,578,484 A (HOROSZEWICZ) 26 November 1996 (26/11/96) see entire document.	2 --- 2
X --- Y	EADES ET AL. Identification and analysis of a class 2 alpha-mannosidase from <i>Aspergillus nidulans</i> . <i>Glycobiology</i> . 1998, Vol. 8, pages 17-33, see entire document.	1-14 --- 1-14
X --- Y	INOUE ET AL. Molecular cloning and nucleotide sequence of the 1,2-alpha-D-mannosidase gene, msds, from <i>Aspergillus saitoi</i> and expression of the gene in yeast cells. <i>Biochimica et Biophysica Acta</i> . 1995, Vol. 1253, pages 141-145, see entire document.	1-14 --- 1-14

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/27210

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/27210

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1, 3-7, 11, 13-14, drawn to a purified mannosidase, polynucleotides and method of making.

Group II, claim 2, drawn to a binding agent that binds the purified protein.

Group III, claims 8-10 and 12, drawn to a method of altering glycosylation pattern of a macromolecule.

The inventions listed as Groups I-III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Mannosidases, their polynucleotides, binding agents and method of altering the glycosylation pattern of macromolecules are well known in the prior art. Thus, the above invention, when considered as a whole does not contribute over the prior art (see Turco et al. J. Biol. Chem., 1982, Vol. 257(15):8674-8679).